REPORT NUMBER 17

ROBUST DETECTION OF FADING NARROW-BAND SIGNALS IN NON-GAUSSIAN NOISE

M. WEISS and S.C. SCHWARTZ

INFORMATION SCIENCES AND SYSTEMS LABORATORY

Department of Electrical Engineering and Computer Science
Princeton University
Princeton, New Jersey 08544

FEBRUARY 1985

Prepared for

OFFICE OF NAVAL RESEARCH (Code 411SP)
Statistics and Probability Branch
Arlington, Virginia 22217
under Contract N00014-81-K0146
SRO(103) Program in Non-Gaussian Signal Processing

S.C. Schwartz, Principal Investigator

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 17	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Robust Detection of Fading Narrow-Band Signals in Non-Gaussian Noise	5. TYPE OF REPORT & PERIOD COVERED: Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): M. Weiss and S.C. Schwartz	8. CONTRACT OR GRANT NUMBER(s): N00014-81-K-0146
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Information Sciences & Systems Laboratory, Department of Electrical Eng. & Computer Sci., Princeton University, Princeton, NJ 08544	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR SRO-103
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research (Code 411SP), Department of the Navy, Arlington, Virginia 22217	12. REPORT DATE: February 1985
	13. NUMBER OF PAGES: 122
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This report will be submitted as part of the Ph.D. Thesis of M. Weiss. Portions of this report will be presented at the IEEE 1985 International Radar Conference, Washington, DC, May 6-9, 1985 and at the IEEE ICASSP, March 26-29, 1985.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

robust detection
non-Gaussian detection
fading narrow-band signals
combined robust estimation-detection
scale invariance
adaptive threshold

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this report, we study procedures for robust detection of slowly fading narrowband signals in nearly Gaussian noise, a common model for radar/sonar systems.

For the classical quadrature matched filter test, we introduce an uncertainty about the envelope statistic and develop a robust test for this one-sample situation. It is demonstrated that this test is capable of protecting only against relatively weak contaminations.

(Continued)

Abstract Continued

With uncertainty taken directly on the noise samples, we develop an estimation-detection theoretic approach: the detection statistic preserves the structure of the quadrature matched filter, but in place of the linear sample mean, a minimax robust estimator of the random amplitude is substituted. This test is shown to be asymptotically maximin optimal (in the sense of Huber) for a wide family of decision rules and for several common target signal models. In addition, the maximin bound on the desired detection probability is sufficiently large to be in the range of practical interest.

This test is then extended to handle the important case of unknown power level. When noise reference samples are available, the extended test utilizes an adaptive threshold, which serves as an estimator of the variance of the amplitude estimator under the noise-only situation. This version is shown to be maximin robust in the class of "Sliding-window" tests. Its implementation in the nominally Gaussian case is relatively simple, as it requires only α-trimmed estimators. When detection is to be performed with the signal-plus-noise observations only, scale invariance is achieved by coupling the "location" estimators with a robust scale estimator. Here, only qualitative robustness is achieved. However, within various parametrized mixture families, the false-alarm (detection) probability can be bounded from above (below).

Finally, Monte Carlo simulation results show that the asymptotic analysis accurately predicts the small sample performance, down to $n = 16$ samples. Moreover, the proposed tests substantially outperform the corresponding robust tests derived from the traditional weak signal assumption.

Robust Detection of Fading Narrow-Band Signals in Non-Gaussian Noise

M. Weiss and S.C. Schwartz
Department of Electrical Engineering
and Computer Science
Princeton University
Princeton, NJ 08544

ABSTRACT

In this report, we study procedures for robust detection of slowly fading narrow-band signals in nearly Gaussian noise, a common model for radar/sonar systems.

For the classical quadrature matched filter test, we introduce an uncertainty about the envelope statistic and develop a robust test for this one-sample situation. It is demonstrated that this test is capable of protecting only against relatively weak contaminations.

With uncertainty taken directly on the noise samples, we develop an estimation-detection theoretic approach: the detection statistic preserves the structure of the quadrature matched filter, but in place of the linear sample mean, a minimax robust estimator of the random amplitude is substituted. This test is shown to be asymptotically maximin optimal (in the sense of Huber) for a wide family of decision rules and for several common target signal models. In addition, the maximin bound on the desired detection probability is sufficiently large to be in the range of practical interest.

This test is then extended to handle the important case of unknown power level. When noise reference samples are available, the extended test utilizes an adaptive threshold, which serves as an estimator of the variance of the amplitude estimator under the noise-only situation. This version is shown to be maximin robust in the class of "Sliding-window" tests. Its implementation in the nominally Gaussian case is relatively simple, as it requires only α-trimmed estimators. When detection is to be performed with the signal-plus- noise observations only, scale invariance is achieved by coupling the

"location" estimators with a robust scale estimator. Here, only qualitative robustness is achieved. However, within various parametrized mixture families, the false-alarm (detection) probability can be bounded from above (below).

Finally, Monte Carlo simulation results show that the asymptotic analysis accurately predicts the small sample performance, down to $n = 16$ samples. Moreover, the proposed tests substantially outperform the corresponding robust tests derived from the traditional weak signal assumption.

Table of Contents

1. INTRODUCTION 1

1.1 Motivation and outline of the work 1
1.2 Review of Huber's results for maximin testing of simple hypotheses 5
1.3 Review and critique of robust tests based on weak signal assumption 8

2.THE NARROWBAND SIGNAL AND NOISE DETECTION ENVIRONMENT 13

2.1 Models and problem statement 13
2.2 Performance degradation of the envelope detector in contaminated Gaussian noise 15

3. ROBUST TEST ON THE SAMPLE ENVELOPE 23

4. ASYMPTOTICALLY MAXIMIN ROBUST TEST 30

4.1 Introduction 30
4.2 Nonrandom signal amplitude 31
4.3 Random signal amplitude 43
4.4 Relationship with weak signal LORD 48

5. ROBUST DETECTORS FOR REALISTIC ENVIRONMENT 51

5.1 Sensitivity to unknown contamination ratio 51
5.2 Extensions of the SSQME test for unknown scale of the nominal density 56
5.3 Unknown signal frequency and implementation complexity 74
5.4 Finite sample simulation results 78

APPENDICES 100

A. Small sample performance of the sign detector 100
B. An approximation for the distribution of the envelope detector 102
C. The maximin test for a single observation 107
D. Statistics of the SSQME test with unmatched frequency 110

REFERENCES 113

1. INTRODUCTION

1.1 Motivation and Outline

The classical theory of signal detection is based on a complete statistical characterization of the signals and interference that are typical of the environment. In many cases, especially for the adverse situations in which radar and sonar systems operate, this precise modeling is either prohibitive or impossible. Consequently, there is a strong motivation to study the design of detection procedures which are robust against deviations from the assumed statistical models.

Following the fundamental work of Huber on robust estimation of location parameters [1] and on robust hypotheses testing [2], extensive applications and further research have appeared in the communication and information sciences literature, of which references [4]-[12] are a representative sample. Inspection of these references reveals that they are mostly concerned with *lowpass deterministic* (i.e., unknown but non-random) signals, except [9] which treats robust detection of *weak stochastic* signals, and [12] where results have recently been presented for *deterministic narrow-band* signals. However, the target signals that are received by radar-sonar systems (which are among the most important engineering applications of statistical detection theory) in typical environments, are represented by more complicated structures than those studied in the above mentioned references. These would include slowly fading narrow-band signals with random amplitude and phase, or range and doppler spread targets [14]. While an extensive detection-estimation theory exists for these complex models when both signal and noise are Gaussian narrow-band processes, c.f [13]- [15], robust detectors have not been devised for these signal structures. The results of [4]-[12] are not directly applicable, except for those of Shin and Kassam in [11] who proposed a rather ad-hoc robust detector for a Rayleigh narrow-band signal.

The main purpose of this work is the design and analysis of robust detection procedures for *coherent* , slowly fading narrow-band signals [1] in nearly Gaussian noise.

[1] This is probably the most important case for modern radar-sonar systems that operate

Under the slow-fading model, the signal random amplitude and phase are essentially constant over the observation period; while their distributions are assumed to be known, the average amplitude level is unknown. Unlike previous works [5],[7]-[10], and [12], where the objective was to robustify the local performance for weak signals, we seek a global solution which will be optimal for any signal amplitude. We achieve this through a different approach in which the detection test statistic is based on a robust estimation of the random amplitude. (A similar structure was proposed by El-Sawy and VandeLinde for the limited case of deterministic lowpass signals). This procedure is shown to be asymptotically maximin optimal in the sense of Huber [2] within a very wide family of decision rules.

Outline Of The Work

Chapter 2 introduces the problem and presents a performance analysis of the common (unrobustified) quadrature detector under situations of contamination. Since in most practical applications even if the noise distribution is known its level is not, the adaptive threshold version of the detector is emphasized. Here, the test statistic is normalized by the Maximum-Likelihood (ML) estimate of the noise level. While this test achieves the desired Constant False Alarm Rate (CFAR) property asymptotically, it is demonstrated that for small sample sizes the false alarm probability (P_{fa}) can increase intolerably under contamination, in addition to a large decrease in the detection probability (P_d).

In Chapter 3 we study a robust test on the "coherent envelope" of the observables, i.e., on the output of a quadrature matched filter which is the sufficient statistic under the purely Gaussian noise case. For this scalar statistic, extension of Huber's [2] robust test is straightforward. It consists essentially in finding the threshold setting and a randomization constant under a least favorable noise distribution. However, it is found that

in severe clutter (correlated noise) background, where coherent filtering ("whitening") is unavoidable.

the proposed test can protect against contaminations that would have roughly doubled the P_{fa} of the unrobustified quadrature receiver. Beyond that, the maximin bound on P_d decreases rapidly as a function of the contamination. Hence, this test is rather inappropriate for radar/sonar, where very low P_{fa} is required, and the percentage contamination (ϵ) can be much larger than this probability.

In Chapter 4 an asymptotically robust test is proposed and analyzed. The common large sample assumption is an unavoidable necessity in order to get explicit functionals for the error probabilities, on which optimization according to a maximin criterion becomes possible. However, we abandon the weak signal assumption that has dominated the literature. By avoiding this assumption, we hope that the proposed test will be adequate for the small sample sizes typical of radar/sonar systems. Hence, the proposed test statistic is derived from minimax robust ML estimation of the amplitude. This is achieved by applying Huber's [1] M-estimator of "location" to the quadratures samples, and then independence of the uniform unknown phase is obtained by summing the squares of these estimates, exactly as in the quadrature matched filter. This test is shown to be maximin robust for sufficiently large values of the desired P_d (in the range of practical interest) and for several common target signal models. Finally, the weak signal locally optimal robust detector for our problem is outlined, and is shown to be a local approximation to the derived M-estimation/detection structure.

In Chapter 5, further uncertainties that must be considered in practical applications are studied. With regard to ϵ in the mixture model [1], it is shown numerically that even a very pessimistic design with ϵ=0.5 incurs a rather small additional loss compared to the case when ϵ is known exactly. More important, the robust detector is extended to handle an unknown scale (power level) of the nominal noise density in the mixture family, by coupling it with a robust estimator of the scale. While it is not possible to exhibit maximin properties of the extended test (as is always the case in multi-parameter problems [3]), it is shown numerically to be qualitatively robust in the sense that an upper

bound on P_{fa} and a lower bound on P_d can be guaranteed within various parametrized mixture families. The latter bound is quite close to the P_d obtained by the optimal detector for the nominal Gaussian noise. When "noise reference" samples are available, a similar estimator-detector with an adaptive threshold, which serves as an estimator of the variance of the amplitude estimator, is shown to be maximin in the class of "sliding window" detectors. Due to its CFAR property for *any* noise distribution, the restriction for sufficiently high P_d of Chapter 4 is relaxed.

Moreover, when the nominal noise is Gaussian, the M-estimators can be replaced by α– trimmed estimators, which are asymptotically equivalent in probability and simpler to compute. The same structure can be generalized for non-Gaussian *nominal* p.d.f. with the appropriate optimal robust L-estimators. (The problems of uncertainty in the scale/variance of the nominal density and in ϵ have not been treated before in detection problems). Finally, the signal frequency which is also unknown in any realistic application (due to doppler shifts), is treated by constructing a bank of contiguous tests which covers the uncertainty range. It is shown that the detectability loss for signals whose frequency straddle between adjacent such "filters", is asymptotically identical with that incurred by the usual periodogram (FFT) detector.

The price that is paid for the improved capability is a substantial increase in signal processing complexity, compared to the linear FFT bank of filter-detector. However, the most demanding nonlinear processing required here consists of several levels of data rankings, and this operation is becoming possible for real-time radar/sonar applications with VLSI and VHSIC technology.

The final section of chapter 5 presents a thorough Monte-Carlo study of the small-sample performance of the various robust detectors. The simulation results show that the performance predicted by the asymptotic analysis of the "sliding-window" robust detectors is essentially maintained even for sample size $= 16$, in guaranteeing a high lower bound on the detection probability as well as in controlling the false alarm-level.

When noise-reference samples are not available, somewhat larger sample sizes are necessary for convergence to the asymptotic prediction. In addition, the corresponding weak-signal locally optimal robust detector is shown to produce high losses for small desired P_{fa} and large deviation of the signal frequency. These losses are attributed to poor convergence to the Gaussian distribution as a result of hard limiting on the test statistic, when the signal is sufficiently large to allow high detection probability. On the other hand, our proposed tests limit the influence of outliers *around* the arbitrary signal amplitude, and converge rapidly to the asymptotic values.

As a consequence of the asymptotic analysis and simulation study, our tentative conclusion is that the proposed robust detectors emerge as quite adequate for successfully treating most of the important uncertainties that are encountered in the real-world radar-sonar detection environment.

1.2 Review of Huber's Results for Maximin Testing of Simple Hypotheses

Huber [2] considered and solved the following problem. Let $\{x_i\}_{i=1}^n$ be a sequence of independent random variables, and let $\{P_0,P_1\}$ be distinct probability measures on the real line with the corresponding densities $\{f_0,f_1\}$ with respect to some measure. Assume that the likelihood ratio (LR) $f_1(x)/f_0(x) = L(x)$ almost surely, where L(x) is a monotone function.

Let **M** be the set of all probability measures on the real line and $0<\epsilon<1$ a given number. The uncertainty in the distribution of the observations is introduced by expanding the *simple* hypothesis P_0 and *simple* alternative P_1 into composite ones by a mixture model-

$$H_0: \quad \mathbf{P}_o = \{Q \in \mathbf{M} \mid Q = (1-\epsilon)P_0 + \epsilon C_0, \quad C_0 \in \mathbf{M}\} \tag{1.1a}$$

$$H_1: \quad \mathbf{P}_1 = \{Q \in \mathbf{M} \mid Q = (1-\epsilon)P_1 + \epsilon C_1, \quad C_1 \in \mathbf{M}\} \tag{1.1b}$$

Actually, Huber's setup is more general, it allows also the following neighborhoods of the nominal model: total variation, Prohorov distance, Kolmogorov distance and L'evy

distance. Moreover, different ϵ_i are allowed in the H_i.

The problem is to find the most robust test in the maximin sense between $\mathbf{P}_0$ and $\mathbf{P}_1$, i.e., to find a saddle-point pair of test $d^*(\mathbf{x})\in\mathbf{D}$, where $\mathbf{D}$ is the class of all decision rules, and distributions $q_i^*\in\mathbf{M}$ such that

$$\sup_{d\in\mathbf{D}} \beta(d,q_1^*) = \beta(d^*,q_1^*) = \inf_{q\in\mathbf{P}_1} \beta(d^*,q) , \tag{1.2}$$

subject to

$$\sup_{\substack{d\in\mathbf{D}\\ q\in\mathbf{P}_0}} \alpha(d,q) = \alpha(d^*,q_0^*) = \alpha_0 \tag{1.3}$$

Here, $\beta(d,q)$ is the power of the test (detection probability) at density q:

$$\beta(d,q) = Prob\{d(\mathbf{x})=H_1 \mid H_1 \text{ is true } (q\in\mathbf{P}_1)\} \tag{1.4}$$

and $\alpha(d,q)$ is the level (false alarm probability)

$$\alpha(d,q) = Prob\{d(\mathbf{x})=H_1 \mid H_0 \text{ is true } (q\in\mathbf{P}_0)\} \tag{1.5}$$

The meaning of the criterion is clear- d^* is the best Neyman-Pearson (NP) decision rule for the lest-favorable pair $\{q_0^*,q_1^*\}$.

As Huber showed, the most robust test is a NP test on the pair

$$q_0^* = \begin{cases} (1-\epsilon)f_0(x) & L(x)<L'' \\ \frac{(1-\epsilon)}{L''}f_1(x) & L(x)\geq L'' \end{cases} \tag{1.6}$$

$$q_1^* = \begin{cases} (1-\epsilon)f_1(x) & L(x)\geq L' \\ (1-\epsilon)L' f_0(x) & L(x)<L' \end{cases} \tag{1.7}$$

The numbers L' and L'' are determined such that q_0^* and q_1^* are legitimate density functions.

For ϵ sufficiently small (a condition which is equivalent to disjointness of $\mathbf{P}_0$ and $\mathbf{P}_1$) the normalizing equations have a unique solution with $0\leq L' \leq L'' <\infty$. The LR between q_0^* and q_1^* is thus given by "soft-limiting" the nominal LR to-

$$l(x;L',L'') = \begin{cases} L' & when \quad L(x)\leq L' \\ L(x) & when \quad L' <L(x)<L'' \\ L'' & when \quad L(x)\geq L'' \end{cases} \tag{1.8}$$

and a maximin robust test is a randomized NP test on $T(\mathbf{x}) = \prod_{i=1}^{n} l(x_i; L', L'')$:

$$d^*(\mathbf{x}) = \begin{cases} H_1 & \text{for} \quad T(\mathbf{x}) > t \\ H_1 \ \textit{with probability} \ c & \text{for} \quad T(\mathbf{x}) = t \\ H_0 & \text{for} \quad T(\mathbf{x}) < t \end{cases} \tag{1.9}$$

The quantities t and c are determined from the right side equality in Eq.(1.3). Note that in general the test must be randomized for arbitrary α_0 since $T(\mathbf{x})$ takes its values at the limiter end-points with finite probability.

It is interesting to point to some peculiarity in the result. The least-favorable densities are constrained by (1.6-1.7), and in particular the contamination C_1 can not be related in any other way to C_0. While this is reasonable in a game situation against an intelligent opponent, it seems unlikely for signal detection problems where the uncertainty in $\mathbf{P}_1$ is induced by that in $\mathbf{P}_0$ and is not affected by the presence or absence of an additive signal whose characteristics are assumed to be known. For example, in a "location shift" problem ($H_0: x_i = n_i$, $H_1: x_i = n_i + a$) we would like $\mathbf{P}_1$ to be the class of all distributions that are shifted to the right by a from those in $\mathbf{P}_0$, and specifically $q_1^*(x) = q_0^*(x-a)$, which *can not* also satisfy (1.6-1.7). Thus, it is suggested that a better solution might exist for this physical formulation of the problem; unfortunately, it is unknown [1]. The asymptotic reformulation in section 4 essentially avoids this peculiarity.

It should be emphasized that Huber's proof of optimum robustness relies heavily on two assumptions which are not valid for the problems that are considered in this work: a) H_0 and H_1 are simple hypotheses- P_0 and P_1 do not include any unknown (relevant or nuisance) parameters. b)the observations are independent r.v.'s.

[1] Private communication with P.J.Huber,12/1983.

1.3 Review and Critique of Robust Tests Based on Weak Signal Assumption

Often in detection problems, the true value of the signal amplitude A is small but otherwise unknown. In such cases it is plausible, mostly as a theoretical nicety which permits an analytical (and usually quite simple for implementation) solution, to employ the locally optimum detector (LOD) structure. This is the detector structure which maximizes the derivative of the power function (detection probability) at A=0, for a given test level (probability of false-alarm), [25]. Under suitable regularity conditions on the densities, the LOD is identical with the detector which is obtained by taking the leading term of a series expansion for the likelihood-ratio in powers of the SNR around zero [26]. It also maximizes Pitman's efficacy [27] which is a suitable weak signal measure of performance, when the number of the observed samples $n \rightarrow \infty$.

The weak signal local optimality criterion was first extended to problems of robust detection by Martin and Schwartz [4], and has been widely applied since then, c.f. [5],[7]-[10],[12]. Instead of seeking a maximin relation on the detection probability as in Eq. (1.2), it was proposed in [4] to design for an asymptotic $(n \rightarrow \infty)$ maximin relation on the slope of the power function at A=0 :

$$\sup_{d \in D} \beta'(d, q_0^*) = \beta'(d^*, q_0^*) = \inf_{q_0 \in \mathbf{P}_0} \beta'(d^*, q_0) \tag{1.10}$$

where $\beta'(d,q) = \frac{d}{dA}\beta(d,q \mid A) \mid_{A=0}$ and subject to the false-alarm constraint of Eq. (1.3). The interpretation is that d^* is the LOD for the least-favorable density q_0^*. This criterion has resulted in all the above mentioned references in a *limiter-correlator* structure: the locally optimal non-linearity of the LOD receiver is robustified against ϵ - contamination by inserting a soft limiter at its output and then correlating the non-linearly transformed observations with the known signal sequence. This will be denoted in the following as LORD - LO Robust Detector.

The performance of LOD schemes is commonly evaluated using asymptotic measures such as the Asymptotic Relative Efficiency (ARE), which requires both the assump-

tions of large n (to satisfy the central limit theorem) and vanishingly weak SNR. In any practical engineering application, the number of samples must be finite [1] and thus the input SNR must be reasonably large to obtain meaningful detection probability. It follows that LOD schemes might perform poorly in practical applications. Several examples of peculiar and very slow convergence of the finite sample RE to the ARE, can be found in [28]-[31]. These references include examples where the finite-sample/large-signal ranking of detector performance are actually different from ARE prediction. Another new example is given in Appendix A. The large SNR performance of the LOD would be extremely poor when the test non-linearity redescends or even vanishes except for some region around the origin, i.e. $l(x)=0 \forall \mid x \mid \geq c$. In these cases we will obviously get $\lim_{A>>c} \beta \rightarrow 0$, in contrast to the desired consistency of the power function with increasing SNR. This situation was actually obtained in [10]. In principle, this undesired property of LOD designs might be corrected by switching between two detectors, where one of them is the LOD and the other one is some amplitude-consistent detector. The switching should occur as a result of a threshold crossing by an estimator of the signal amplitude $\hat{A}$. This estimator must obviously be robust against deviations in the assumed noise model to prevent incorrect switching. While this heuristic proposal has not been analyzed , it does suggest that an optimal robust detector should be based on a robust amplitude estimator, as will be studied in section 4.

Inherently, the LORD scheme is subject to the same consistency problem. In addition, investigation of the previously mentioned references reveals that they all exhibit one or more of the following peculiarities and shortcomings:

[1] In radar-sonar systems, n is directly related to the total search time of the desired sector, to the desired maximum non-ambiguous detection range, and to the spatial resolution of targets. Moreover, the detectability of *coherent* signals is (asymptotically) governed by the average integrated $SNR = nA^2/2\sigma^2$. Since transmitters are usually constrained by the average rather than peak power, it is only the product nA^2 that matters and improved detectability can be obtained by increasing A as well as n. Hence, in view of the other system design goals that were mentioned in the beginning of this note, n is usually in the range 1-100, and even in 1-3 samples for very long range systems, in contrast with the common theoretical assumptions.

i) The support of the contaminating density c in Eq. (1.1) is restricted to lie exclusively on the exterior of some interval $[-x_1,x_1]$, c.f. [9]. This might not be a severe limitation since the greatest deterioration in performance of the conventional detectors occurs for contaminations that are in the far tails of the nominal density.

ii) The nominal density itself *is not* a member of the mixture family P_0 in which a least-favorable density is sought, but rather "close" to it in some measure, c.f. [12].

iii) The test is maximin robust only for $\alpha \geq \alpha(\epsilon)$, which is roughly in the range 0.05-0.2, c.f. [4],[5], and [7]. Hence, no robustness is guaranteed for the most important range of small α: 10^{-4}–10^{-8} !

It is clear that the maximin relation (1.10) is neither sufficient nor necessary for obtaining the desired maximin solution in terms of the detection probability itself. Since Eqs. (1.2) and (1.3) are contradictory in nature when A can take any value in some interval, the most that can be expected from the LORD approach is as shown in Fig. 1.1: the shaded region on the β axis indicates the α's obtained for the optimal detector for any density q_0 in the mixture family; the slope of any $\beta(q_0)$ at A=0 agrees with (1.10) and (1.2) is satisfied for all signal amplitudes that are greater than some critical value : $A \geq A_c$ or equivalently $\beta \geq \beta_c$.

However, the power function of a LORD could as well behave as shown in Fig. 1.2 where (1.3) and (1.10) are satisfied, but the curve of $\beta(q_0^*)$ dominates that of $\beta(q_0)$ $\forall$ $A \geq 0$ and $\forall q_0 \in \mathbf{P}_0$. (Other situations are also possible). Nevertheless, if $\forall q_0 \in \mathbf{P}_0$ the distance between the curves in Fig. 1.2 could be made sufficiently small, i.e.- $|\beta(d^*,q_0)-\beta(d^*,q_0^*)| \leq \delta(\epsilon,\alpha_0,A) \rightarrow 0$, the detector would be practically robust, though not in the most general maximin sense. (In that case, q_0^* actually becomes the most-favorable density for the detection probability !)

Closer study of the structure of all the LORD tests in the above mentioned references reveal that implicit in all of them is a "robust estimator of zero amplitude" -i.e.,

the test is actually based on an approximation of the leading term in the expansion of an appropriate version of Huber's robust estimator in a power series around zero amplitude. This observation establishes another motivation for the asymptotically robust test of section 4 which utilizes robust estimation of the amplitude without any small signal approximation, and thus essentially avoids all of the above mentioned problems.

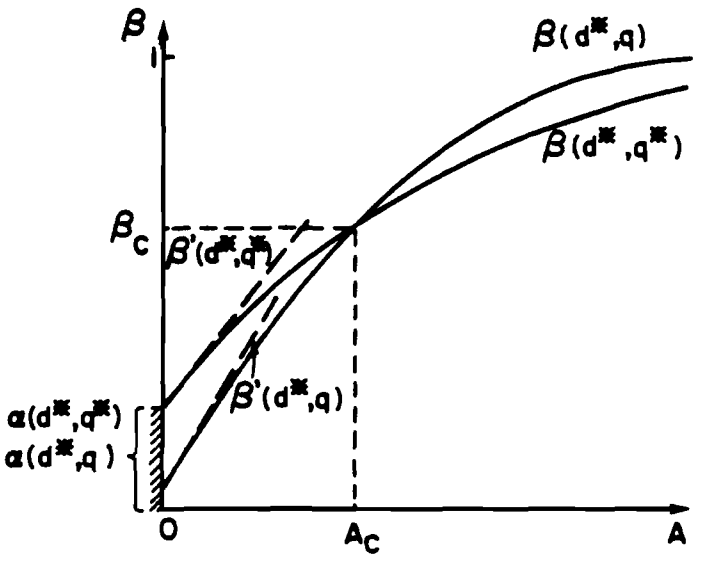


Fig. 1.1 Qualitative power function curves for a desired maximin robust detector of signal with unknown amplitude.

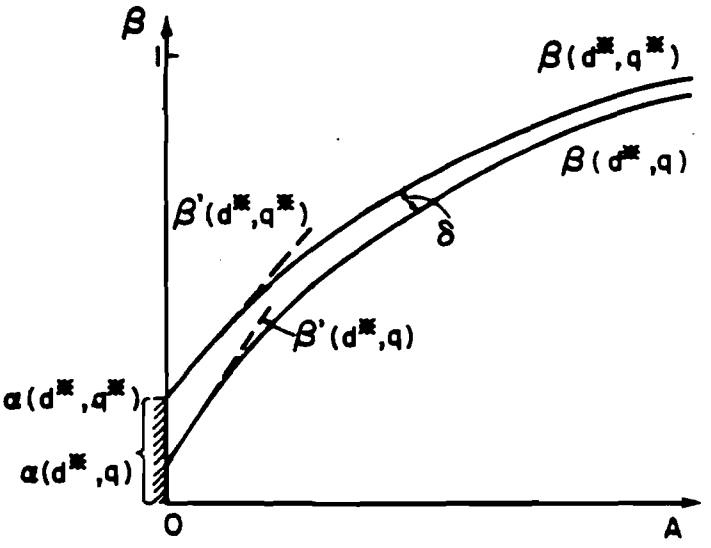


Fig. 1.2 Qualitative power function curves for a locally-optimal robust detector which is not maximin with respect to the detection probability.

2. THE NARROW-BAND SIGNAL AND NOISE DETECTION ENVIRONMENT

2.1 Models and Problem Statement

The problem to be considered is that of robust detection of a slowly fading narrow-band signal with unknown phase and amplitude in nearly Gaussian narrow-band noise. This signal model is typical of most coherent radar targets [13] and also appears in some sonar problems and also in communication over tropospheric links [14].

Specifically, the discrete signal plus noise samples that are observed at the systems input terminals are

$$z_i = As_i \cos(\omega t_i + \phi) + n_i \qquad ,i = 1,...,n \tag{2.1}$$

The n_i are narrow-band noise samples around the *known*[1] center frequency ω :

$$n_i = n_{c_i} \cos\omega t - n_{s_i} \sin\omega t \tag{2.2}$$

$\{\mathbf{n_c},\mathbf{n_s}\}$ are assumed to be zero mean, independent [2] and identically distributed (i.i.d) random variables (r.v.), from a nominally Gaussian ϵ -contaminated mixture :

$$f\,(\mathbf{n_c},\mathbf{n_s}) = \prod_{i=1}^{n} f_0(n_{c_i}) f_0(n_{s_i}) \tag{2.3}$$

where

$$f_0(n_{c_i} = x) = f_0(n_{s_i} = x) = (1-\epsilon)\frac{1}{\sqrt{2\pi}\sigma}\exp(\frac{-x^2}{2\sigma^2}) + \epsilon c\,(x) \qquad ,c \in \mathbf{M} \tag{2.3a}$$

ϕ is the unknown signal phase which is common for all samples (coherent detection) and uniformly distributed over $[0,2\pi]$. It does not convey any information on the target and is thus a nuisance parameter which has to be averaged out. $\{s_i\}_{i=1}^{n}$ is a sequence of known, positive, finite-energy amplitude modulations [3]. The amplitude A which is also

[1] The case of unknown doppler shift will be treated later by the usual "bank of filters" approach.

[2] Notice that sample -to -sample independence either excludes the common situation of detection in correlated clutter, or it corresponds to a sub-optimal scheme where the observables z_i are obtained by pre- whitening the input sequence.

[3] Known phase modulations can be introduced without much complication.

constant for all the observed samples will be considered in the sequel as either deterministic but unknown or random with some known distribution of unknown scale. [4]

It is well known, c.f. [13], that in the purely Gaussian noise case ($\epsilon = 0$), a uniformly most powerful (UMP) test, independent of the amplitude being deterministic or random with an arbitrary distribution, is equivalent to a threshold test on the "coherent sample envelope" $R(\mathbf{x},\mathbf{y})$:

$$R(\mathbf{x},\mathbf{y}) = (\sum_{i=1}^{n} x_i s_i)^2 + (\sum_{i=1}^{n} y_i s_i)^2 \tag{2.4}$$

where the in-phase and quadrature samples are obtained by applying the input to a pair of lowpass mixers:

$$x_i = 2z_i \cos\omega t \mid_{LPF} = As_i \cos\phi + n_{c_i} \tag{2.5}$$

$$y_i = 2z_i \sin\omega t \mid_{LPF} = As_i \sin\phi + n_{s_i} \tag{2.6}$$

Since the quadrature components are obtained by a reversible operation, they are equivalent statistics for the problem, which can thus be reformulated as:

$$H_o: \quad \begin{cases} x_i = n_{c_i} \\ y_i = n_{s_i} \end{cases} \tag{2.7a}$$

$$H_1: \quad \begin{cases} x_i = n_{c_i} + As_i \cos\phi \\ y_i = n_{s_i} + As_i \sin\phi \end{cases} \tag{2.7b}$$

It is desired to obtain a maximin relation like Eq. (1.2) for all $A \geq A_c$ (recall the discussion in section 1.3) subject to the false-alarm constraint Eq. (1.3).

Direct application of Huber's [1] solution for this problem is inappropriate since the samples under H_1 are actually dependent due to the common random phase ϕ, even when A is deterministic and known. However, a "naive" extension of Huber's test is pos-

[4] These slow fading assumptions correspond to common radar targets whose cross section (RCS) fluctuates with a correlation time much longer than the "blip" duration nT, T= intersample period, but much shorter than the scan-to- scan period. This is a most frequent situation [13].

sible if one considers ϕ as a non-random but unknown and adopts an "estimation-detection" procedure. A set of $M \rightarrow \infty$ tests is performed, where the j^{th} test is matched to $\phi_j = j/2\pi M$ and its objective is a decision about the presence of a signal *of phase* ϕ_j. Each of the individual tests is then of the *correlator-limiter* type found in [4], but the limiter break-points are now functions of the s_i and also of ϕ_j. Since phase information is irrelevant, the final decision can be reached by accepting H_1 if at least one of the individual tests has accepted it.

We have not studied how many parallel tests have to be performed to achieve acceptable performance with arbitrary signal phases between the adjacent "filters". It is conjectured that it might be prohibitive, especially when the doppler shift is also unknown so that a two-dimensional "filter-bank" (phase, frequency) has to be constructed.

In the absence of any optimal procedure of design for the finite sample problem, we suggest to weaken somewhat the requirements. We will do that in section 3 by assuming that the observable is only the envelope statistic of Eq. (2.4) rather than the original quadrature samples. We can then construct a most robust test against the uncertainty in the distribution of that *scalar* statistic.

In section 4, we will extend the asymptotic estimation-detection approach of El-Sawy and VandeLinde [6] to our problem, utilizing some of Huber's [1] results on robust estimation of a location parameter. It is shown that if A is replaced by a robust minimax estimate, the resulting test statistic has a limited robust property.

2.2 Performance Degradation of the Conventional Envelope Detector.

In this section, the degradation in performance of the envelope detector of Eq. (2.4), (which is the UMP detector for slow fading narrow-band signals in narrow-band Gaussian noise) will be analyzed when it operates in a background which comes from a Huber-Tukey mixture family. The decrease in detection probability as well as the increase in false alarm probability are of interest, where the latter is of an extreme

importance in an automated search system. These results will serve as a basis for comparison with the performance of the robust detectors that are developed later in this work.

Radar and sonar systems usually operate in a highly non-stationary and non-homogeneous noise background. Except for the receiver thermal noise which is stationary and Gaussian, the statistical properties of interference that is the outcome of external sources, (such as clutter reflections from ground sea or aerial objects in the radar environment, or ambient noise and reverberations in the sonar case [14], [22]), rapidly change in time and space. Therefore, even when the Gaussian assumption is adequate, the fixed threshold detector based on the envelope statistics of Eq. (2.4) is of little value, since its actual false-alarm rate will fluctuate intolerably according to the changes in the background noise level. A common adaptive detector for these situations, which is invariant to the power level of the background noise, compares the envelope statistics from the "test-cell" with an adaptive threshold derived from a "noise-reference" channel. Specifically, $\{\mathbf{X},\mathbf{Y}\}$ represent the narrow-band observation samples in the hypothesis-testing cell, $\{\mathbf{U},\mathbf{V}\}$ are the noise-reference samples which are assumed to have *the same distribution* as that of $\{\mathbf{X},\mathbf{Y}\}$ under H_0. The adaptive test is:

$$d(\mathbf{x},\mathbf{y},\mathbf{u}.\mathbf{v})=H_1 \qquad if\ f \qquad R(\mathbf{x},\mathbf{y})\geq tW(\mathbf{u},\mathbf{v}) \tag{2.8}$$

where

$$R=I^2+Q^2\ ; \quad I(\mathbf{x})=\frac{1}{n}\sum_{i=1}^{n}x_i\ ; \quad Q(\mathbf{y})=\frac{1}{n}\sum_{1}^{n}y_i \tag{2.9}$$

$$W(\mathbf{u},\mathbf{v})=\frac{1}{M}\sum_{j=1}^{M}(U_j^2+V_j^2) \tag{2.10}$$

$$U_j=\frac{1}{n}\sum_{i=1}^{n}u_{ij} \quad ;V_j=\frac{1}{n}\sum_{i=1}^{n}v_{ij} \tag{2.11}$$

The threshold multiplier t is determined as to achieve the desired false -alarm rate. (Notice that there are M reference vectors for the test vector, all of dimension n).

In search systems where a target presence is to be detected in some spatial sector, the noise reference samples are easily obtainable by applying the same signal processing

of the hypothesis testing channel to adjacent resolution cells in range, doppler or bearing coordinates, c.f. [18]-[20]. This detector is known as the Cell- Averaging CFAR(CA-CFAR) , Mean Level Detector(MLD), or Sliding Window Detector (SWD). When the number M of reference samples grow asymptotically, its performance approaches that of the UMP detector for the fixed-variance Gaussian noise. Its structure is also almost identical with that of a narrow-band version which can be derived from the Two-samples(i.e., noise-reference) Students-t detector [21], which has some optimum properties among unbiased and invariant tests for detection in Gaussian noise of unknown level. (In the extended t-test, the noise variance is estimated from both the test and reference samples in contrast to (2.8). However, the difference in performance is quite small when M is sufficiently large, and the structure of Eq. (2.8) enables "sliding-window" (SW) operation where R as well as $U_j^2+V_j^2$ are generated sequentially *by the same hardware* during the search).

In the following we assume that the background interference has an epsilon-contamination mixture density, where the nominal as well as the contaminating densities in Eq.(2.3) are Gaussian: $f = (1-\epsilon)N(0,1)+\epsilon N(0,c^2)$. The performance of the detector is discussed first when the number of samples n is asymptotically large, and secondly for the finite sample case.

a) n $\rightarrow\infty$

When n is asymptotically large, all the quantities I, Q, U_j, and V_j of Eqs.(2.9)-(2.11) are Gaussian random variables. Hence, it is obvious that this detector is asymptotically nonparametric under the null hypothesis; i.e., it is CFAR for any probability density of the input test and reference samples for which the central limit theorem holds.

Under the alternative, for a Rayleigh distributed slow-fading signal the detection-probability of this detector is uniquely determined by the integrated Signal-to-Noise ratio (c.f. [18], [20]): $SNR = nA^2/2\sigma^2$. Since $\sigma^2=(1-\epsilon)+\epsilon c^2$, it is clear that even a small amount of high power(c^2) contamination can reduce the effective SNR by orders of

magnitude, thus spoiling detectability completely.

b) Finite n

With regard to the detection probability a similar behavior as above is to be expected, as it is not significantly affected by the tails of the distribution. However, it is no longer true that the detector remains CFAR for moderate and small sample sizes. Since in practice radar-sonar systems are designed for very low false alarm probabilities, it is the tails of the distribution of the test statistics that counts, and even a slight deviation from Gaussianity of the input samples will suffice to substantially increase the area under the tails.

In Appendix B, an approximation is derived for the distribution of the coherent envelope R, which approaches the desired one-sided exponential distribution for large n. [1] The approximation is based on an expansion of the distribution in a series of orthonormal Laguerre polynomials, where its leading term is (with proper normalization such that E(R)=1) the Gamma density

$$f(R)=\frac{\alpha^{\alpha} e^{-\alpha R} R^{\alpha-1}}{\Gamma(\alpha)} \quad , \quad R \geq 0 \tag{2.12}$$

where $\alpha=(1+k/2n)^{-1}$ and k is the kurtosis of the input samples. This is given for a Gauss-Gauss mixture by

$$k = \frac{E(x^4)}{E^2(x^2)} - 3 = 3\epsilon(1-\epsilon)\left[\frac{c^2-1}{1-\epsilon+\epsilon c^2}\right]^2 \tag{2.13}$$

Notice that when $\epsilon \rightarrow 0$ but $\epsilon c^2 >> 1$, $k \rightarrow 3/\epsilon >> 1$. Table 2.1 demonstrates that k can be very far away from zero for a Gauss-Gauss mixture, thus α is also much smaller than 1 leading to higher false-alarm probability.

[1] Actually, an exact analytic expression was derived for the special case of a Gauss-Gauss mixture, but it was found to be very difficult to compute numerically.

ϵ c	3.162	10	31.62
0.001	.238/.992	24.3/.568	748.5/.041
0.02	2.02/.941	73.5/.303	245.4/.115
0.1	6.06/.841	22.3/.589	26.47/.547

Table 2.1. Kurtosis(left entry) and α (right) , Gauss-Gauss mixture, n=16.

Figures 2.1 thru 2.5 depicts the false alarm probability vs. the threshold multiplier t, where t is set according to the Gaussian assumption. In each figure, the graphs marked "known sigma" corresponds to the adaptive threshold detector of Eq.(2.8) with M $\rightarrow\infty$; in that case the variance estimation is error-free and the detector is identical in performance to the UMP detector for the known variance case. The graphs marked " ϵ=0" give P_{fa} when the input is indeed Gaussian. In all figures, n=16 and M=8. This is sufficiently small for poor convergence to Gaussianity of the sample means in the tails, but is representative of the number of samples actually employed in most systems. The adaptive-threshold curve with ϵ=0 was computed from $P_{fa}=(1+t/M)^{-M}$, c.f. [18], [20]. The Laguerre approximation was computed only for the M $\rightarrow\infty$ case, since for the adaptive threshold test no similar approximation is possible. [2] The squares were obtained from a Monte-Carlo simulation; it is observed from the figures that the Laguerre approximation is quite reasonable and agrees closely with the simulation results in its region of low variance(roughly the reciprocal of the number of runs in the simulation). The main conclusion from the figures is that the increase in P_{fa} is very high, for desired values of 10^{-4} and lower whenever $\alpha<0.9$, and that this deterioration is ordered in correspondence with the deviation of k and α from their values at the Gaussian density. This increase in P_{fa} is much more severe for the adaptive-threshold detector with finite number of reference cells; in Fig. 2.5, for example, the threshold setting for 10^{-6} at the Gaussian produces about $3\cdot10^{-2}$ at the contaminated mixture.

[2] Since it is not possible to obtain analytic expressions for the moments of the ratio R/W of Eq. (2.8). While it is possible to derive a similar approximation to the distribution of W and then to integrate numerically Pr{R>tW}, this approach was not taken.

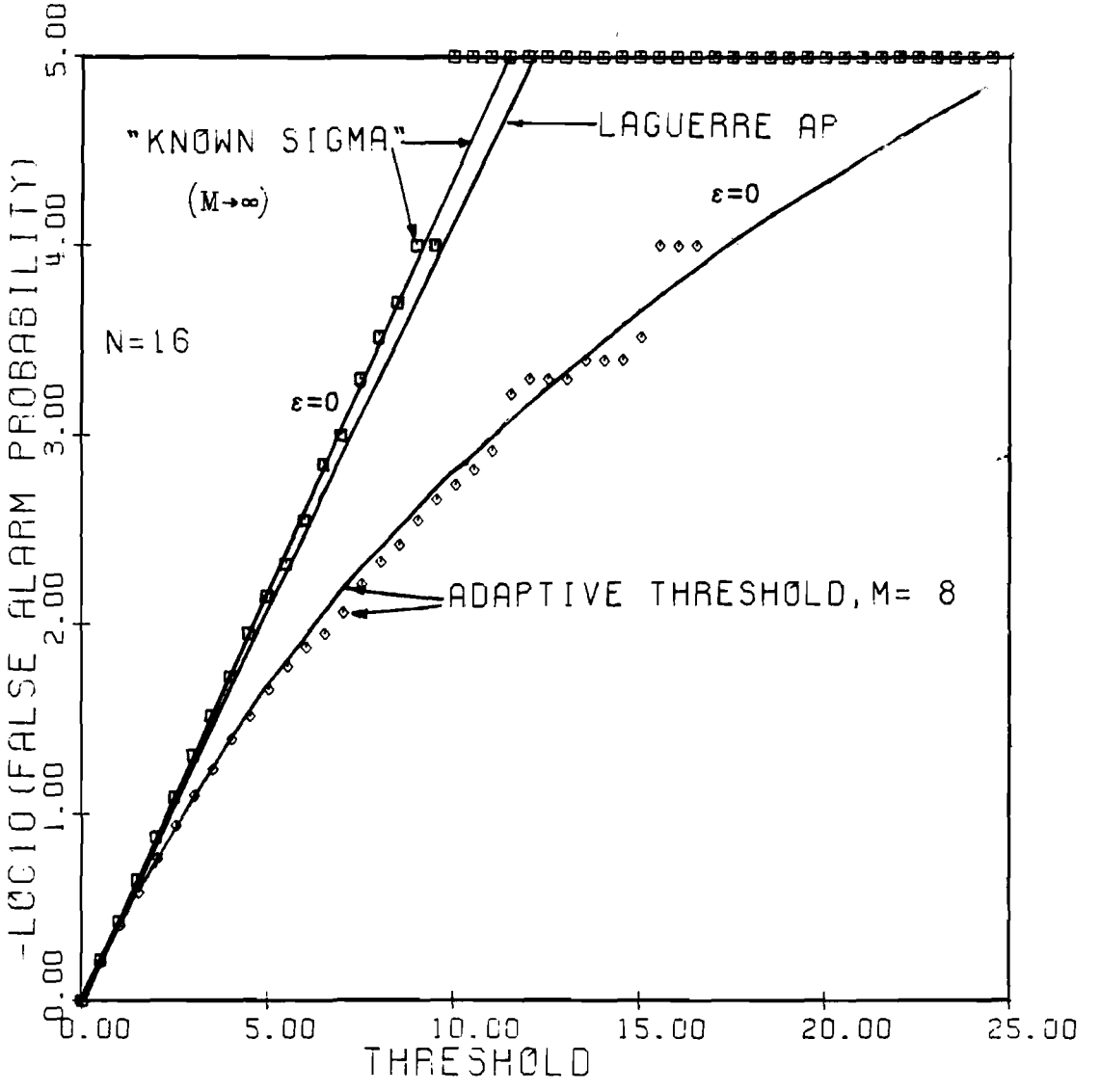


Fig. 2.1 false-alarm probability vs. threshold multiplier. Normal-Normal mixture, ε=0.001,c=3.162,k=.2384,α=.9926,10,000 monte-carlo runs.

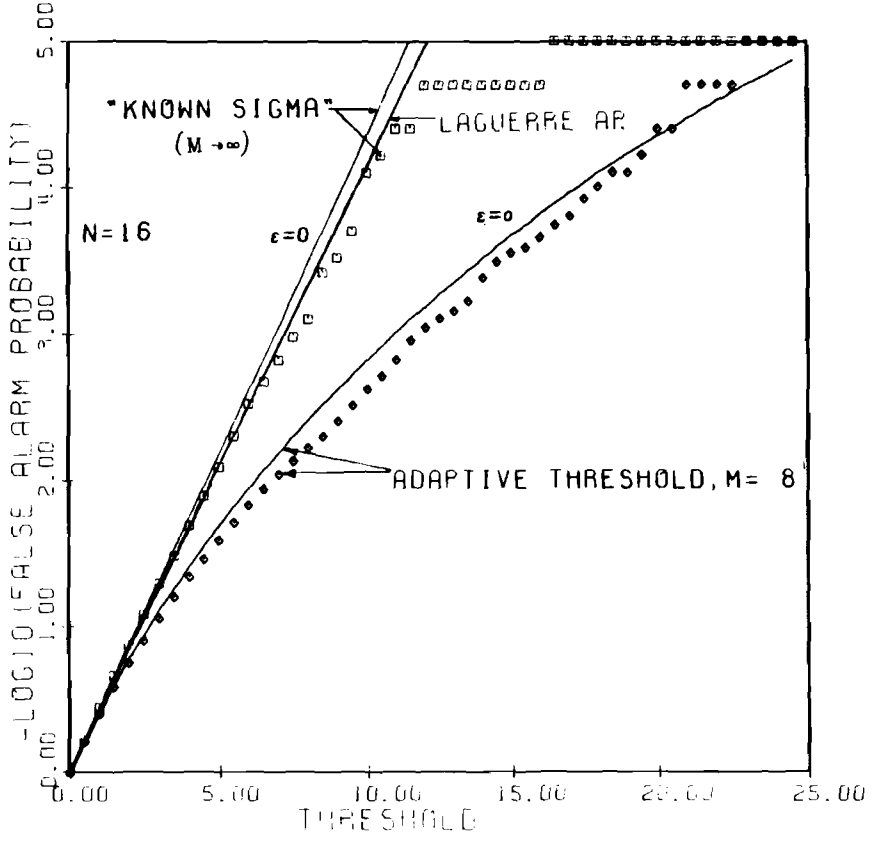


Fig. 2.2 false-alarm probability vs. threshold multiplier. Normal-Normal mixture, ε=0.01,c=3.162,k=2.025,α=.9405,50,000 monte-carlo runs

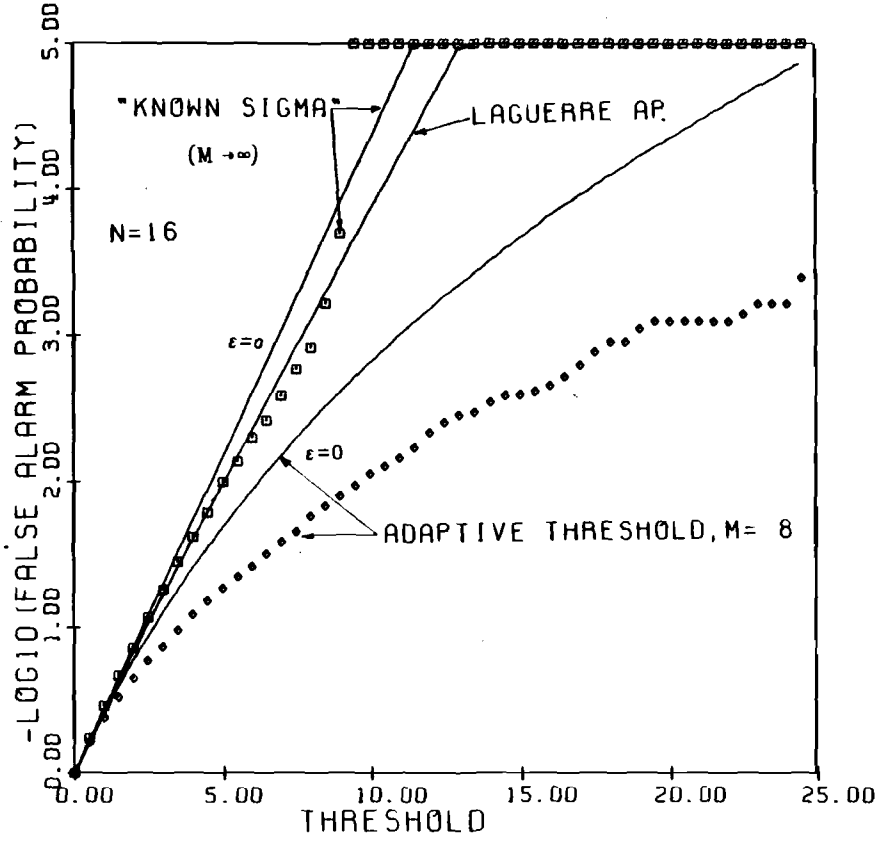


Fig. 2.3 false-alarm probability vs. threshold multiplier. Normal-Normal mixture, ε=0.1,c=3.162,k=6.06,α=.8408,10,000 monte-carlo runs.

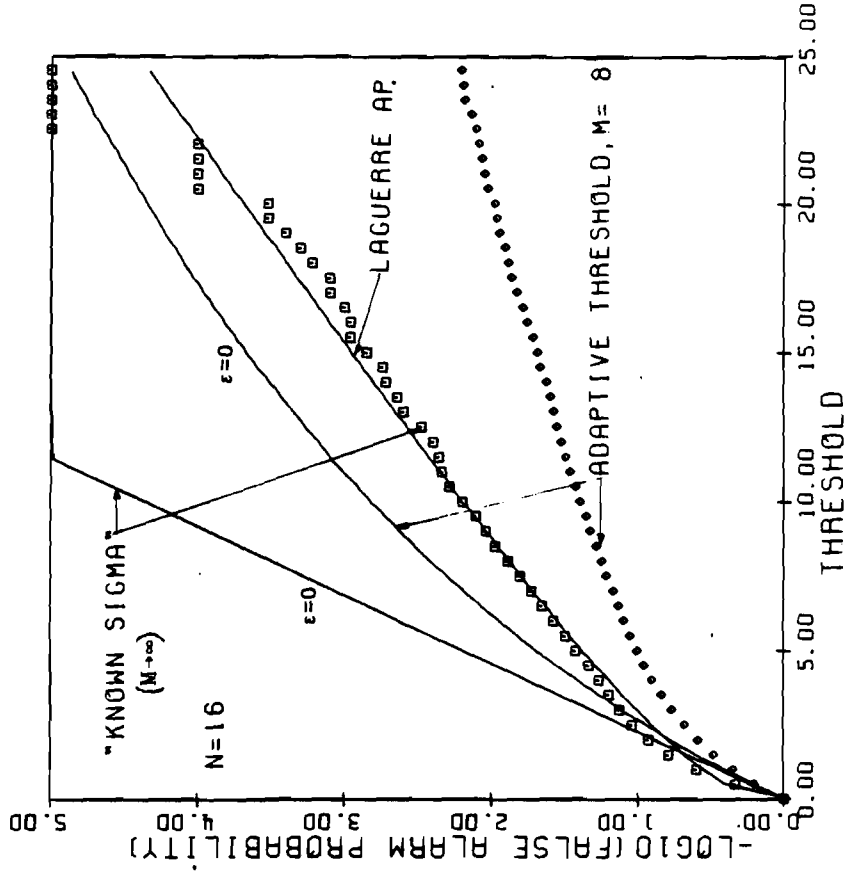


Fig. 2.5 false-alarm probability vs. threshold multiplier. Normal-Normal mixture, ε=0.01,c=10.,k=73.51,α=.3033,:0,000 monte-carlo runs.

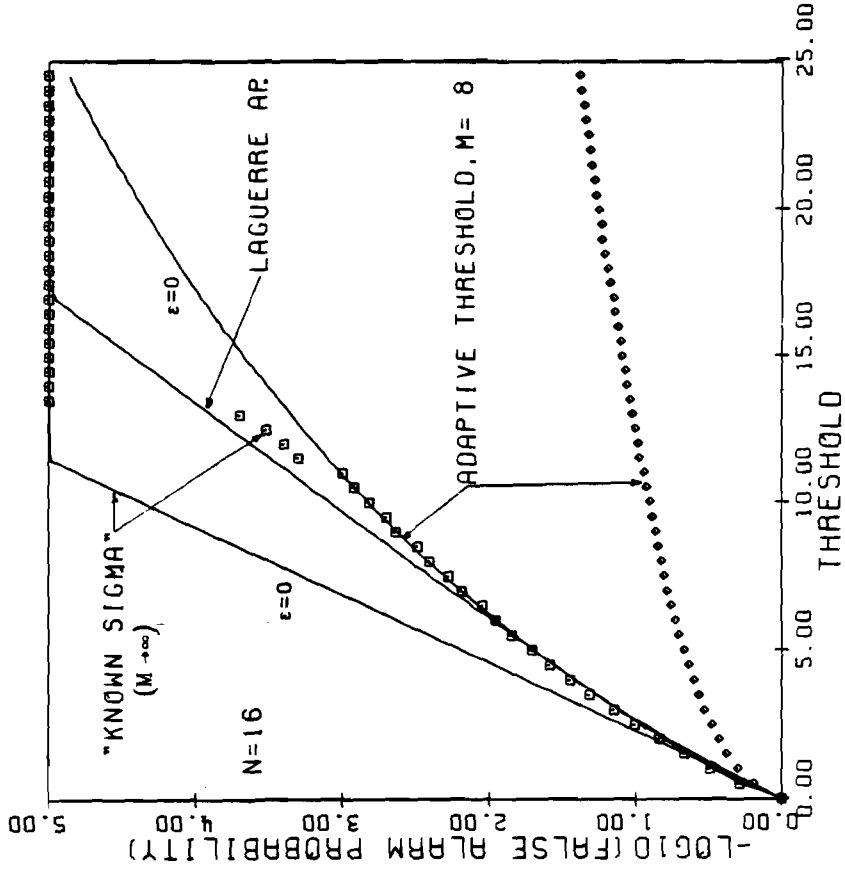


Fig. 2.4 false-alarm probability vs. threshold multiplier. Normal-Normal mixture, ε=0.1,c=10.,k=22.27,α=.5896,:0,000 monte-carlo runs.

3. ROBUST TEST ON THE SAMPLE ENVELOPE

When non-contaminated hypotheses contain nuisance (vector) parameters $\boldsymbol{\theta}_i$ with *known* distribution they are actually simple hypotheses, and the optimal test is a Neyman-Pearson test on the likelihood ratio of the averaged densities:

$$L(\mathbf{x}) = E_{\theta_1}[f_1(\mathbf{x} \mid \boldsymbol{\theta}_1)]/E_{\theta_0}[f_0(\mathbf{x} \mid \boldsymbol{\theta}_0)] \tag{3.1}$$

An attempt to incorporate averaging into the proof of optimality of Huber's test fails due to the fact that the marginal density is in general not a product of densities of the components x_i: $E_\theta[f(\mathbf{x} \mid \theta)] \neq \prod_{i=1}^{n} E_\theta[f(x_i \mid \theta)]$, unless $n = 1$. When $n = 1$, the structure of Huber's test extends to the case where there are nuisance parameters; the only correction needed is to replace $L(x)$ of section 1.2 by $\bar{L}(x) = E_\theta[f_1(x \mid \theta)]/f_0(x)$, provided it is a monotone function of x.

Therefore, in the absence of an optimal procedure for the design of a robust test on the original observables $\{x_i, y_i\}_{i=1}^n$ of Eq. (2.7), it is proposed to weaken the requirements and to consider instead a most robust test on some *scalar* statistic. Since for $\epsilon = 0$ the sample envelope $R(\mathbf{x}, \mathbf{y}) = (\sum x_i)^2 + (\sum y_i)^2$ is a sufficient statistic for the problem, and its distribution is independent of the phase, it is natural to try to robustify a test based on R.

Thus we will consider the modified problem:

$$\mathbf{P}_0 = \{q(R) \in \mathbf{M} \mid q(R) = (1 - \epsilon')\, f_0(R) + \epsilon'\, c_0(R),\ c_0 \in \mathbf{M}\}$$
$$\mathbf{P}_1 = \{q(R) \in \mathbf{M} \mid q(R) = (1 - \epsilon')\, f_1(R) + \epsilon'\, c_1(R),\ c_1 \in \mathbf{M}\}$$

where ϵ' is the amount of contamination on the envelope R.[1]

[1] If one knows how to find ϵ_i of the original observables (by physical reasoning, estimation or guessing), he should be able to apply his procedure for finding ϵ'. Nevertheless, it can be verified that if the uncertainty in the measure is given by a mixture model, non-linear operations on random variables do not change ϵ and summation of n i.i.d. contaminated r.v's results in $\epsilon_n = 1 - (1 - \epsilon_i)^n$, from which we get $\epsilon' = 1 - (1 - \epsilon_i)^{2n} \approx 2n\,\epsilon_i$.

We now apply this approach to Gaussian contaminated quadrature noise components and Rayleigh distributed signal amplitude.[2] It is well known [13] that the nominal densities are given by:

$$f_1(R \mid A)=\sigma_R^{-2}\exp[-\sigma_R^{-2}(R+n^2A^2)]I_0(A\sqrt{R}/\sigma^2)\ ,\ R\geq 0 \tag{3.2}$$

$$f_1(R)=E_A[f_1(R \mid A)]=[\sigma_R^2(1+S_n)]^{-1}\exp[-R/\sigma_R^2(1+S_n)]\ ,\ R\geq 0 \tag{3.3}$$

where $\sigma^2=E_{f_0}(n_c^2)=E_{f_0}(n_s^2)$, $\sigma_R^2=2n\sigma^2, S=A_0^2/2\sigma^2$ is the input signal to noise ratio under the nominal conditions, $A_0^2=E(A^2)$ and $S_n=nS$.

The least favorable pair is given by Eqs.(1.6-1.7), with $f_1(R)$ from (3.3) and $f_0(R)$ by the same expression with $S_n=0$, and

$$\bar{L}(R)=(1+S_n)^{-1}\exp(KR)\ ,\ R\geq 0,\quad K\triangleq S_n/\sigma_R^2(1+S_n) \tag{3.4}$$

As $\bar{L}(R)$ is monotone, an equivalent test statistic is

$$T(R)=\begin{cases}\sigma_R^2 r' & ,\quad R\leq\sigma_R^2 r'\\ R & ,\quad \sigma_R^2 r'<R\leq\sigma_R^2 r''\\ \sigma_R^2 r'' & ,\quad R>\sigma_R^2 r''\end{cases} \tag{3.5}$$

where the limiter breakpoints are found by solving the normalization equations (C.2-C.3) of Appendix C. These equations yield

$$\exp[-r'(1+S_n)^{-1}]+(1+S_n)^{-1}\exp[r' S_n(1+S_n)^{-1}][1-\exp(-r')]=(1-\epsilon')^{-1} \tag{3.6}$$

$$1-\exp(-r'')+(1+S_n)\exp[-r'' S_n(1+S_n)^{-1}]\exp[-r''(1+S_n)^{-1}]=(1-\epsilon')^{-1} \tag{3.7}$$

It is shown in Appendix C that the resulting maximin test can take three different forms, depending on ϵ and the desired α_0. Roughly, when ϵ/α_0 is large, it is a randomized test where H_1 is decided with probability $c<1$ if $R>t$. For intermidiate values of ϵ/α_0, it is a deterministic threshold test on R, and for small values, H_0 is decided with probability c, if $R<t$. Equation (3.7) is explicitly solved by

[2] This correspons to Swerling I target model [17], which is a very good description of the radar cross section (RCS) fluctuations for microwave frequencies, where the target size is much larger than the wavelength so that the quadrature components are due to summation of many independent reflectors and the central limit theorem holds - [13].

$$r'' = \ln[S_n(1-\epsilon')/\epsilon'] \tag{3.8}$$

while solutions for the first one can be found numerically from the set of solutions $\{a < 1\}$ of

$$S_n a + a^{-S_n} = (1+S_n)(1-\epsilon')^{-1} \quad , \quad a = \exp[-r'(1+S_n)^{-1}] \tag{3.9}$$

Table 3.1 shows the values of the "sufficiently small" ϵ_c' for which $\mathbf{P}_0$ and $\mathbf{P}_1$ are disjoint and the maximin test exists ($L'(\epsilon') \leq L''(\epsilon'), \forall \epsilon' \leq \epsilon_c'$). These are given as a function of the integrated signal to noise ratio S_n, the natural measure of distance between the hypotheses for this problem.

S_n	10^{-2}	10^{-1}	1	10	10^2	10^3	∞
ϵ_c'	$3.5 \cdot 10^{-3}$	$4 \cdot 10^{-2}$	.2	.42	.49	.498	.5

Table 3.1 Values of critical contamination ϵ_c' vs. integrated nominal signal-to-noise ratio $S_n = nE(A^2)/2\sigma^2$.

A suitable measure of performance of a robust test is how far is the lower bound on the power $\beta(d^*, q^*)$, from the power of the Neyman-Pearson test for $\epsilon = 0$, β_0. For our problem $\beta_0 = \alpha_0^{1/(1+S_n)}$, and using (3.7) with (C.7) and (C.15) of Appendix C, we get

$$\beta(d^*, q^*) = \begin{cases} (1-\epsilon')^{S_n/(1+S_n)} (\alpha_0 - \epsilon')^{1/(1+S_n)} & , \quad \epsilon' \leq \dfrac{S_n \alpha_0}{1+S_n} \\ \dfrac{\alpha_0}{1+S_n} \left[S_n \dfrac{1-\epsilon'}{\epsilon'} \right]^{S_n/(1+S_n)} & , \quad \epsilon' > \dfrac{S_n \alpha_0}{1+S_n} \end{cases} \tag{3.10}$$

The different expressions correspond to cases b) and c) in the appendix. The last one is valid for those desired values of α_0 when the limiter is "effective" and one must resort to a randomized test, which in turn causes faster decrease of the power curve. Case a) of the appendix is never applied in this Rayleigh signal example, as it was found numeri-

cally that it corresponds to false-alarm probabilities higher than 0.5.

Figures 3.1 - 3.2 depict $\beta(d^*,q^*)$ versus ϵ' for different values of ϵ' and S_n. All the graphs are clearly characterized by a sharp "knee" at $\epsilon' \approx \alpha_0 S_n/(1+S_n) \approx \alpha_0$, such that when ϵ' is larger than this critical value the power of the test deteriorates rapidly. Notice that even an increase of orders of magnitude in the effective signal to noise ratio S_n does not help to alleviate the problem. The figures also show the false alarm probability of the optimal detector for the *uncontaminated case*, when it actually operates in a worst-case contamination under H_0: $\alpha(N.P.,w.c.)=\alpha_0(1-\epsilon')+\epsilon'$. (Under H_1 the worst case power is $(1-\epsilon')\beta_0$ which is insignificantly lower than β_0 for the small ϵ' s considered).

Thus we conclude that the proposed test can protect against contaminations which would have roughly doubled the false alarm probability of the Neyman-Pearson test ($\epsilon' \lesssim 2\alpha_0$). Beyond that, this protection is achieved at an intolerable price of decrease of the power. In radar applications, α_0 is very small and its variation is significant only for an order of magnitude changes. Thus the test proposed in this section is not very useful in many practical situations.

A somewhat similar robust detector for a Rayleigh signal in nearly Gaussian noise was recently studied by Shin and Kassam [11]. Motivated by insight gained from the structure of various robust detectors, they suggested inserting a limiter-squarer nonlinearity at the outputs of the in-phase $I=\sum x_i$ and quadrature $Q=\sum y_i$ channels, summing the outputs and comparing to a threshold. The amount of limiting was optimized numerically and the performance was analyzed. Comparing Fig. 6 of [11] with Figs. (3.1)-(3.2) here, it is observed that the scheme of [11] is somewhat better. However, the numerical analysis of [11] was carried out *only* for the range $\epsilon \leq \alpha_0$, and it is not known whether the performance degrades further when ϵ is orders of magnitude higher than α_0.[3] At the least, this comparison demonstrates that it is advisable to "push back the

[3] Private communication with the authors.

limiting" (or more precisely, an adequate censoring of the outliers as will be clear from Chapter 4) as far as possible from the threshold comparison point to the original observables.

Figure 3.1 a) False alarm probability that would have resulted when *worst-case* contamination is applied to the UMP test for purely Gaussian noise, vs. ϵ. b) Maximin bound on the detection probability vs. ϵ. $\alpha_0 = 10^{-4}$

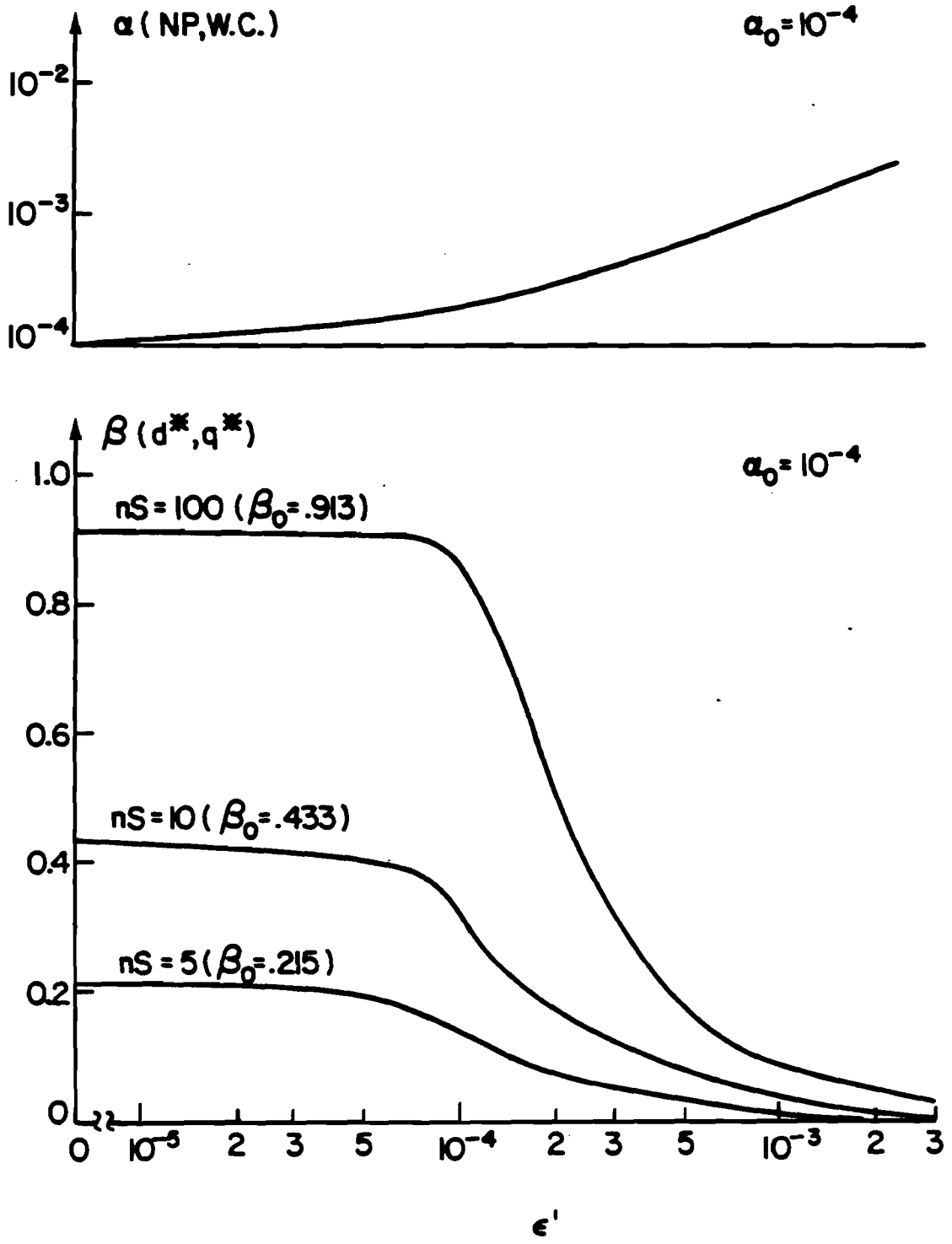
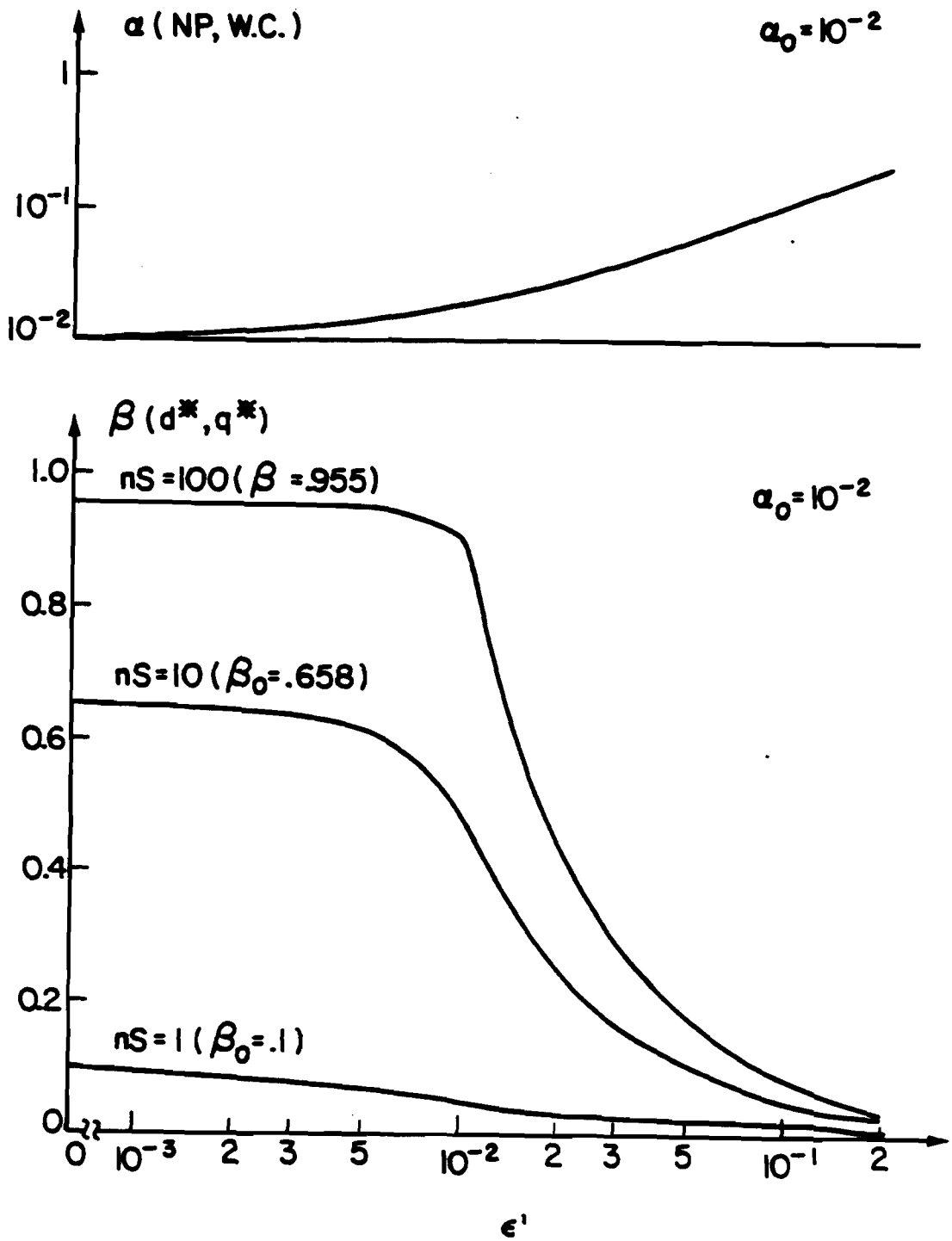


Figure 3.2 a) False alarm probability that would have resulted when *worst-case* contamination is applied to the UMP test for purely Gaussian noise, vs. ϵ. b) Maximin bound on the detection probability vs. ϵ. $\alpha_0 = 10^{-2}$



4. ASYMPTOTICALLY MAXIMIN ROBUST TEST

4.1 Introduction

In this section a test will be derived which is asymptotically ($n \rightarrow \infty$) most robust in a maximin sense. In a following section, we will show by means of simulation that this test maintains its performance even for small sample sizes.

Utilizing some of Huber's [1] results on robust estimation of a location parameter, El-Sawy and VandeLinde [6] derived an asymptotically most robust test for the problem of detecting a completely known signal in additive noise of uncertain distribution. A similar approach is possible (and turns out to be optimal) for the problem formulated in Section 2.1. We observe the similarity between the narrowband slowly fading coherent signal and the lowpass deterministic signal: for a *given* received sample $\{\mathbf{X}, \mathbf{Y}\}$, $A\cos\phi$ and $A\sin\phi$ are essentially unknown location parameters for the quadrature signal components. Moreover, in purely Gaussian noise the unknown phase (which is a nuisance parameter) is "averaged out" in the UMP test statistic $R = I^2 + Q^2$, where I and Q are the sample-means of the in-phase and quadrature components of the narrowband observations. I and Q are the Maximum-Likelihood estimates of the component locations and thus R is a good estimate of the amplitude (squared) which is the "true" parameter of the problem, i.e., the one that distinguishes between H_0 and H_1.

Building further on insight gained from the UMP statistic R, we observe that in the purely Gaussian case the detection probability $\beta = E_A E_\phi Prob\{R > t \mid A, \phi\}$ is analytically tractable as a result of the unbiasedness and Gaussianity of I and Q under both hypotheses. Since in order to exhibit a maximin property for a composite testing problem, it is necessary to obtain analytical expressions for the error probabilities functionals, it appears natural to preserve the structure of the R test, but to replace the sample-mean estimates with robust estimates of the locations $\{A\cos\phi, A\sin\phi\}$, which will have the following properties. The estimates should be:

i) Shift invariant (or unbiased) for all n.

ii) Efficient and consistent.

iii) Asymptotically Gaussian.

iv) Possess the minimax property on the estimation variance (which is translated into the effective SNR when iii is valid).

v) Maintain the above properties for medium and even small n.

Huber's M-estimator [1], and the α-trimmed mean estimator (when the nominal p.d.f. is Gaussian) have all these properties. This facilitates the proof of maximin optimality of the proposed test when the amplitude is non-random. For random amplitudes which are of more interest in practical applications, the expectation of the detection probability over the amplitude must be taken. It turns out that for the most frequently used target model in radar applications (Rayleigh amplitude or Swerling case I), the maximin property of the proposed test is only "almost" preserved (recall Fig. 1.2), but it is maintained exactly for a higher order chi-squared amplitude model. (Rayleigh is first order while constant amplitude is the limit of the chi-square family when the number of degrees of freedom tends to infinity).

For the sake of continuity and convenience in the exposition, we will begin with the non-random amplitude case, and then move to the more realistic random amplitude models. Throughout this section the amount of contamination ϵ in the Huber-Tukey model, as well as the variance of the nominal p.d.f., are assumed to be known. These strong assumptions which are inherent in previous work on robust detection [4]-[12], but are rarely satisfied in practice, will be relaxed in Section 5.

4.2 Non-random Signal Amplitude

Before presenting the test, we need some definitions. Define a class of functions Ψ on R^1 such that $\{L\,,\,l\} \in \Psi$ if:

(4.1)

1) L is convex, symmetric about the origin and strictly increasing for positive argument.

2) $l(x) = dL/dx$ is continuous.

3) for all $q \in \mathbf{P}_0$, $0 < E_q[l^2(x)] < \infty$.

4) for all $q \in \mathbf{P}_0$, $\frac{\partial}{\partial\theta} E_q[l(x-\theta)]$ exists and is nonzero at $\theta = 0$.

Define the estimates by the following implicit equations

$$\hat{A}_x(l) = Arg\{\sum_{i=1}^{n} l(x_i - S_i\hat{A}_x) = 0\} \quad , \quad \hat{A}_y(l) = Arg\{\sum_{i=1}^{n} l(y_i - S_i\hat{A}_y) = 0\} \tag{4.2}$$

($\hat{A}_x$ and $\hat{A}_y$ are the robust M-type estimates [1] for the location parameters $A\cos\phi$ and $A\sin\phi$ of the observables in Eq.(2.7b), for a given realization of the random variables A and ϕ.) Define:

$$T_n(l_0) = [\hat{A}_x(l_0)]^2 + [\hat{A}_y(l_0)]^2 \tag{4.3}$$

where

$$l_0(x) = -d[\log q_0^*(x)]/dx \tag{4.4}$$

and $q_0^* \in \mathbf{P}_0$ minimizes Fisher's information:

$$I(q_0^*) \leq I(q) = \int [\frac{d}{dx} \log q(x)]^2 q(x)dx \, , \quad \forall\, q \in \mathbf{P}_0 \tag{4.5}$$

Consider a threshold test on $T_n(l_0)$

$$d_n^*(\mathbf{x}, \mathbf{y}) = \begin{cases} H_1 \, , & T_n(l_0) \geq t_n(l_0) = t(l_0)/n \\ H_0 \, , & \text{otherwise} \end{cases} \tag{4.6}$$

and let $T(l)$ be any other test based on $\hat{A}_x(l)$ and $\hat{A}_y(l)$ with $l \in \Psi$ but $L \neq -log\ q_0^*$. The following proposition states the asymptotic maximin properties of the test d_n^*, which we will subsequently refer to as the SSQME (Sum of SQuared M-type Estimates) test.

Proposition 4.1 Consider the detection problem defined on the observables of Eq. (2.7) with the uncertainty in the independent quadrature noise samples obeying an $\epsilon-$ mixture model $\mathbf{P}_0$ with symmetrical p.d.f. When A is nonrandom and ϕ is uniformly distributed

on $[0, 2\pi]$, among the class of all tests $T(l)$ based on $\hat{A}_{\mathbf{x}}(l)$ and $\hat{A}_{\mathbf{y}}(l)$ of Eq. (4.2) with $l \in \Psi$, there exists a least-favorable p.d.f. $q_0^* \in \mathbf{P}_0$ and a constant $0 \leq \beta_c \leq 1$ such that the SSQME test d_∞^* defined in (4.6) with $n \to \infty$ achieves:

$$\sup_{d \in T(l)} \beta(d, q_0^*) = \beta(d_\infty^*, q_0^*) = \inf_{q \in \mathbf{P}_0} \beta(d_\infty^*, q) \tag{4.7}$$

whenever $\beta(d_\infty^*, q_0^*) \geq \beta_c$ and subject to

$$\sup_{\substack{d \in T(l) \\ q \in \mathbf{P}_0}} \alpha(d, q) = \alpha(d_\infty^*, q_0^*) = \alpha_0 \tag{4.8}$$

Moreover, when the nominal p.d.f. in $\mathbf{P}_0$ is Gaussian, the nonlinearity l_0 is a symmetric soft limiter whose breakpoints depend on the noise variance but not on the signal amplitude:

$$l_0(x/\sigma^2, -K/\sigma, K/\sigma) = \begin{cases} K(\epsilon)/\sigma & , \quad x \geq K(\epsilon)\sigma \\ x/\sigma^2 & , \quad -K(\epsilon)\sigma \leq x < K(\epsilon)\sigma \\ -K(\epsilon)/\sigma & , \quad x < -K(\epsilon)\sigma \end{cases} \tag{4.9}$$

with $K(\epsilon)$ and ϵ related by

$$f_0(K(\epsilon))/K(\epsilon) - \Phi(-K(\epsilon)) = \epsilon/2(1-\epsilon) \tag{4.10}$$

and $f_0(x) = d\Phi(x)/dx$ is the standard normal p.d.f.

Comments

1) Note that the theorem shows the saddle-point pair of p.d.f. and test only among a certain class $T(l)$. We would like the left relation in (4.7) to be valid for *any* hypotheses test on the original observables $\{\mathbf{X}, \mathbf{Y}\}$. Unfortunately, this could not be proved, unlike the case of a deterministic signal in [6], as will be clear in the sequel. However, the class $T(l)$ is quite large. As limiting cases it contains the sample means of $\mathbf{X}$ and $\mathbf{Y}$ when $\epsilon = 0$[1], the sample medians ($\epsilon = 1$) and with

[1] Thus the SSQME test is better (at least asymptotically) than the "scalar" envelope test which was derived in the preceding section.

weak regularity conditions which usually holds [1] $\hat{A}_x$ and $\hat{A}_y$ can be any translation invariant statistics, such as $\hat{A}_x = \sum_{i=1}^{n} a_i x_{(i)}$ provided $\sum_{i=1}^{n} a_i = 1$ ($x_{(i)}$ are the rank ordered samples $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$), and similarly for $\hat{A}_y$.

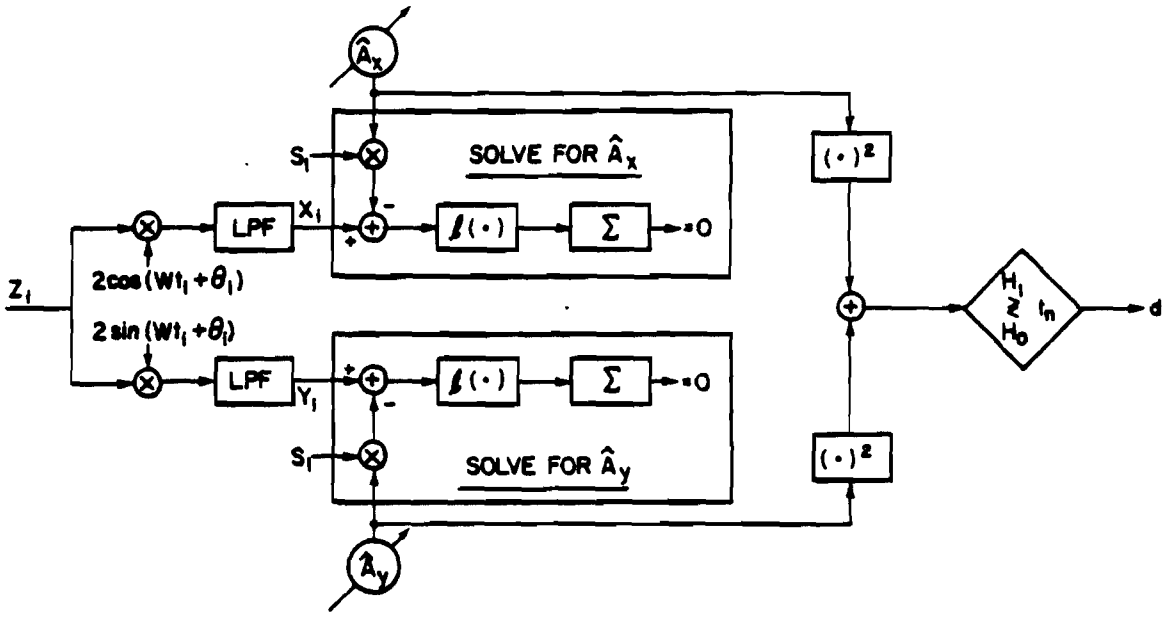
2) The theorem remains valid (but l_0 given by a different expression than (4.9)) for any other description of the uncertainty in the noise, provided the asymptotic variance of the M-estimates (4.2) computed with l_0 is the maximum achievable over the class of densities in $\mathbf{P}_0$. An example is the P family [6]. Others can be found in [1], [3].

3) In contrast to Huber's finite sample size test, this asymptotic test satisfies a more desired and natural optimality criterion. Namely, the maximin relation for the power Eq. (4.7) is specified in terms of the uncertainty in $\mathbf{P}_0$ *alone,* allowing the uncertainty in $\mathbf{P}_1$ to be induced by that in $\mathbf{P}_0$ as a shift of the densities of the quadrature samples (recall the discussion in Section 1.2).

4) Extension to unknown frequency is straightforward by constructing a parallel bank of SSQME tests, the i^{th} best being matched to $w_i = i/nT$ where T is the sampling period. This will be discussed in more detail in section 5.3. The individual test is shown pictorially in Fig. 4.1. Note that a routine for solving for a zero of a function is needed.[2] Thus, a real time implementation in a radar or sonar system calls for substantially larger number of calculations than encountered in the more common detection systems that incorporate linear filtering, FFT and at most, passing the samples through a nonlinearity.

[2] This is true except for the case where the nominal p.d.f. of the uncertainty family is Gaussian, where a considerable simplification is possible as will be shown later.

Figure 4.1- Block-diagram structure of SSQME test for robust detection of slow-fading narrowband signal.



Proof of proposition 4.1 We need to study the asymptotic distribution of $T(l)$, $l \in \Psi$. It can be derived using the following lemmas from [1], with a slight modification from [6].

Lemma 1 Whenever $l \in \Psi$ and q is any symmetric p.d.f., conditioned on A and ϕ, $\sqrt{n}\,(\hat{A}_x(l)-A_x)$ and $\sqrt{n}\,(\hat{A}_y(l)-A_y)$ are asymptotically distributed as normal r.v.'s with zero mean and variance $V(l, q)$

$$V(l, q) = \frac{E_q[l^2(x)]}{C\left[\frac{\partial}{\partial\theta} E_q[l(x-\theta)] \mid_{\theta=0}\right]^2}, \quad C \triangleq \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{i=1}^{n} S_i\right\}^2 \qquad (4.11)$$

where the parameters (A_x, A_y) are the true values of the locations (0 under H_0, $A_x = A\cos\phi$ and $A_y = A\sin\phi$ under H_1).

Lemma 2 Let $q \in \mathbf{P}_0$ be an ϵ-contaminated mixture family where the nominal p.d.f. f_0 is symmetric and twice continuously differentiable, such that $-log\ f_0$ is convex on the convex support of F_0. Whenever $l \in \Psi$, there exists $q_0^* \in \mathbf{P}_0$ which minimizes Fisher's information (4.5) over $\mathbf{P}_0$ and an $l_0 \in \Psi$ given by (4.4). Thus, the asymptotic variance of the M-estimates computed with l_0 at the least favorable p.d.f. q_0^* satisfies the minimax relation

$$\sup_{q \in \mathbf{P}_0} V(l_0, q) = V(l_0, q_0^*) = 1/I(q_0^*) = \inf_{l \in \Psi} V(l, q_0^*) \tag{4.12}$$

where q_0^* is given by

$$q_0^*(x) = (1-\epsilon)\cdot\begin{cases} f_0(x_0)e^{K(x-x_0)} & x \le x_0 \\ f_0(x) & x_0 < x < x_1 \\ f_0(x_1)e^{-K(x-x_1)} & x \ge x_1 \end{cases} \tag{4.13}$$

and $x_0 < x_1$ are the endpoints of the interval where $|f_0'/f_0| \le K$, and K is related to ϵ through

$$\int_{x_0}^{x_1} f_0(x)dx + \frac{f_0(x_0) + f_0(x_1)}{K} = \frac{1}{1-\epsilon} \tag{4.14}$$

We note that the right inequality in (4.12) is a consequence of $\hat{A}_x(l_0)$ and $\hat{A}_y(l_0)$ being the maximum likelihood (ML) estimates of the corresponding location parameters, when the underlying p.d.f. is q_0^*, under *both* hypotheses. In addition, they are efficient under the assumed conditions on Ψ, i.e., they achieve the Cramer-Rao lower bound $I^{-1}(q_0^*)$. The left inequality follows from the fact that q_0^* minimizes $I(q)$ over $\mathbf{P}_0$. Relaxation of the symmetry requirement and other noise uncertainty models are treated in [1], [3] and [6].

From Lemma 1 and the mutual independence of $\mathbf{X}$ and $\mathbf{Y}$, $\hat{A}_x(l)$ and $\hat{A}_y(l)$ are independent and jointly Gaussian when ϕ is fixed. Averaging over ϕ is straightforward with transformation to polar coordinates and yields

$$f\,(\hat{A}_x(l),\,\hat{A}_y(l)\mid A\,) = E_\phi[f\,(\hat{A}_x(l),\,\hat{A}_y(l)\mid A\,,\,\phi)f\,(\phi)] \;=\cdots \tag{4.15}$$

$$= \frac{n}{2\pi V(l\,,\,q\,)}\exp\left[-\frac{T(l)+A^2}{2V(l\,,\,q\,)/n}\right] I_0\left[\frac{A\;T^{1/2}(l)}{V(l\,,\,q\,)/n}\right] \quad,\quad T(l)\;\triangleq\;\hat{A}_x^{\,2}(l)+\hat{A}_y^{\,2}(l)$$

where $I_0(\cdot)$ is the zeroth order modified Bessel function. Also,

$$f\,(T(l)\mid A\,) = \pi f\,(\hat{A}_x(l),\,\hat{A}_y(l)\mid A\,) \tag{4.16}$$

Now, ***restricting*** ourselves to the class of tests based on $\hat{A}_x(l)$ and $\hat{A}_y(l)$ as the new observables of the problem, the likelihood-ratio is $\Lambda(\hat{A}_x(l),\,\hat{A}_y(l)\mid A\,)=\exp(-nA^2/2V)\;I_0[nAT^{1/2}(l)/V\,]$ which is monotone increasing in $T(l)$ for any A, deterministic or random. Therefore, in view of the Neyman-Pearson Lemma $T(l)$ is a sufficient statistic and a threshold test based on it is UMP for H_1: $A>0$ vs. H_0: $A=0$.

Unlike the known lowpass signal case that was treated in [6], it can not be proved here that $T(l)$ is a sufficient statistic of the original observables $\{\mathbf{X},\mathbf{Y}\}$. There, where the additive signal was non-random, use was made of theorems by Wald [23] and Chernoff [24] that a test based on the ML estimate $\hat{\theta}$ (i.e., solution to $\sum l\,(x_i-\hat{\theta})=0$ with $l=(-\log f\,)'\in\Psi$) is asymptotically equivalent to a NP test. In our problem, $T(l_0)$ *is not* the ML estimate of A^2 when the underlying quadrature noise p.d.f.'s are q_0^* [3]. However, its expected value is $E(T(l))=A^2+\frac{2}{n}\,V(l\,,\,q\,)$. Thus, it is asymptotically unbiased and the variance of $T(l_0)$ has a minimax property by virtue of Eq.(4.12). Indeed, it can be shown that

$$Var\,(n^{1/2}T(l)) = 4A^2V(l\,,\,q\,)[1+\frac{A^2}{n}\,V(l\,,\,q\,)]\underset{n\;>>\;A^2}{\longrightarrow}4A^2V(l\,,\,q\,)$$

Hence, $T^{1/2}(l_0)$ is a good robust estimate of A, and as was discussed before, the class

---

[3] The MLE of A is given by solving $\frac{d}{dA}\log E_\phi[\prod_{i=1}^{n} q_0^*(x_i-A\cos\phi)\;q_0^*(y_i-A\sin\phi)]=0$, which is analytically intractable unless q_0^* is identically Gaussian.

$\{T(l)\}$ is quite large.

Having shown that in the restricted class of tests the optimal procedure must be a threshold test on the envelope $T(l)$, it remains to validate Eq.(4.7) subject to (4.8). Using (4.16) with the definition of Marcum's Q-function $Q(a, b) = \int_b^\infty x \exp[-\frac{1}{2}(x^2 + a^2)]\, I_0(ax)dx$, and recalling that when $l \neq l_0$ we are free to adjust the threshold $t(l)$ in (4.6) such that $\alpha(l, q_0^*) = \alpha_0$ while for the pair (l_0, q) the threshold remains $t(l_0)$, the various false-alarm and detection probabilities are as summarized in Table (4.1).

	$\alpha(l, q)$	$\beta(l, q) = Q[a = A\sqrt{n/V(l, q)},\ b = \sqrt{t(l)/V(l, q)}]$
(l, q_0^*)	$\exp[-t(l)/2V(l, q_0^*)]$	$a = A\sqrt{n/V(l, q_0^*)},\ b = b_0$
(l_0, q_0^*)	$\exp[-t(l_0)/2V(l_0, q_0^*)]$	$a = A\sqrt{n/V(l_0, q_0^*)} \triangleq a_0,\ b = \sqrt{-2\log\alpha_0} \triangleq b_0$
(l_0, q)	$\exp[-t(l_0)/2V(l_0, q)]$	$a = A\sqrt{n/V(l_0, q)} = a_0\sqrt{V(l_0, q_0^*)/V(l_0, q)}$ $b = \sqrt{t(l_0)/V(l_0, q)} = b_0\sqrt{v(l_0, q_0^*)/V(l_0, q)}$

Table 4.1- False alarm and detection probabilities for various test-p.d.f. pairs.

Note that the parameter $a^2/2$ defined by

$$\frac{a^2}{2}(l, q) = \frac{nA^2}{2V(l, q)} \tag{4.16a}$$

is the "effective integrated SNR," by analogy with the UMP detector for the purely Gaussian noise where the integrated SNR is given by the same expression with $V \to \sigma^2$.

From the first and second row of the table, by monotonicity of the Q-function with respect to the first argument (a) when the second is fixed, and since from Lemma 2 $\inf_{l \in \Psi} V(l, q_0^*) = V(l_0, q_0^*)$, we get:

$$\sup_{l \in \Psi} \beta(l, q_0^*) = \beta(l_0, q_0^*)$$

This proves the left hand side of (4.7).

From the first column of the second and third row of the table, since the exponent function is monotonic increasing when V increases, by virtue of the other inequality in Lemma 2, we get $\sup_{q \in \mathbf{P}_0} \alpha(l_0, q) = \alpha(l_0, q_0^*)$. It remains to find under what conditions $\beta(l_0, q) = Q(a_0 C, b_0 C) \geq \beta(l_0, q_0^*)$, where $C(q, q_0^*) \triangleq \sqrt{V(l_0, q_0^*)/V(l_0, q)} \geq 1$ for all $q \in \mathbf{P}_0$. Note that the Q function is monotonic in either of its arguments only when the second is fixed, but here they both change proportionally and the inequality is analytically intractable. Since numerical computation of the Q function is notoriously difficult, we give only a sufficient condition by solving for a_0 with a fixed b_0 those values which satisfy $d\beta(l_0, q)/dC \geq 0$, $C \geq 1$. Using known properties of the partial derivatives [15, Appendix F]:

$$\frac{\partial Q(x,y)}{\partial x} = y \exp[-\frac{x^2+y^2}{2}] I_1(xy) \quad , \quad \frac{\partial Q(x,y)}{\partial y} = -y \exp[-\frac{x^2+y^2}{2}] I_0(xy) \qquad (4.17)$$

we obtain an equivalent relation which is sufficient for the right hand side of Eq. (4.7)

$$\frac{I_1(a_0 b_0 C^2)}{I_0(a_0 b_0 C^2)} \geq \frac{b_0}{a_0} \quad , \quad \forall \quad C \geq 1 \qquad (4.18)$$

From the definition of the modified Bessel functions, I_1/I_0 is monotonic increasing from 0 to 1 when the argument increases from 0 to infinity, so it suffices to solve (4.18) for $C = 1$. By continuity and monotonicity of both sides of Eq.(4.18), a solution exists, and upon fixing the false alarm probability ($b_0 = \sqrt{-2 \log \alpha_0}$) and by virtue of the monotonicity of the Q-function with respect to a_0, the solution of Eq. (4.18) with equality defines a critical value β_c such that when $\beta(l_0, q_0^*) = Q(a_0, b_0) \geq \beta_C$, $\inf_{q \in \mathbf{P}_0} \beta(l_0, q) = \beta(l_0, q_0^*) \geq \beta_C$. This completes the proof of proposition 4.1

Some numerical values of β_C are given in Table 4.2. Recall that they are found from a sufficient condition so that they are a conservative estimate for the range of

maximin optimality; we see that for typical false alarm levels $\beta_C \approx 0.60$.

α_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	$\rightarrow 0$
a_0	2.41	3.20	3.86	4.41	4.92	6.16	$\rightarrow b_0$
$\beta_C = Q[a_0, b_0]$	0.68	0.64	0.62	0.60	0.60	0.59	$Q[b_0, b_0]$

Table 4.2 - Critical values of $\beta(l_0, q_0^*)$ for which the SSQME test is maximin robust, non-random amplitude.

We note that the test in [6] for a completely known signal is asymptotically maximin robust only for $\beta_C = 0.5$ (regardless of α_0), a fact that was not directly stated there. Closer examination of the situation in [6] reveals that $\beta_C = 0.5$ is due to the symmetry of the normal p.d.f. of the test statistics around its mean under *both* H_0 and H_1. This is not the case for the SSQME test, and results in a higher value. Nevertheless, when designing a radar, sonar or communication system, the desired goal is in high quality detection. Consequently, systems are specified to achieve β=0.8–0.95, and the technical parameters which result in the required SNR are designed accordingly. Therefore, obtaining the robustness of the SSQME test only for a limited range of β is of no practical limitation.

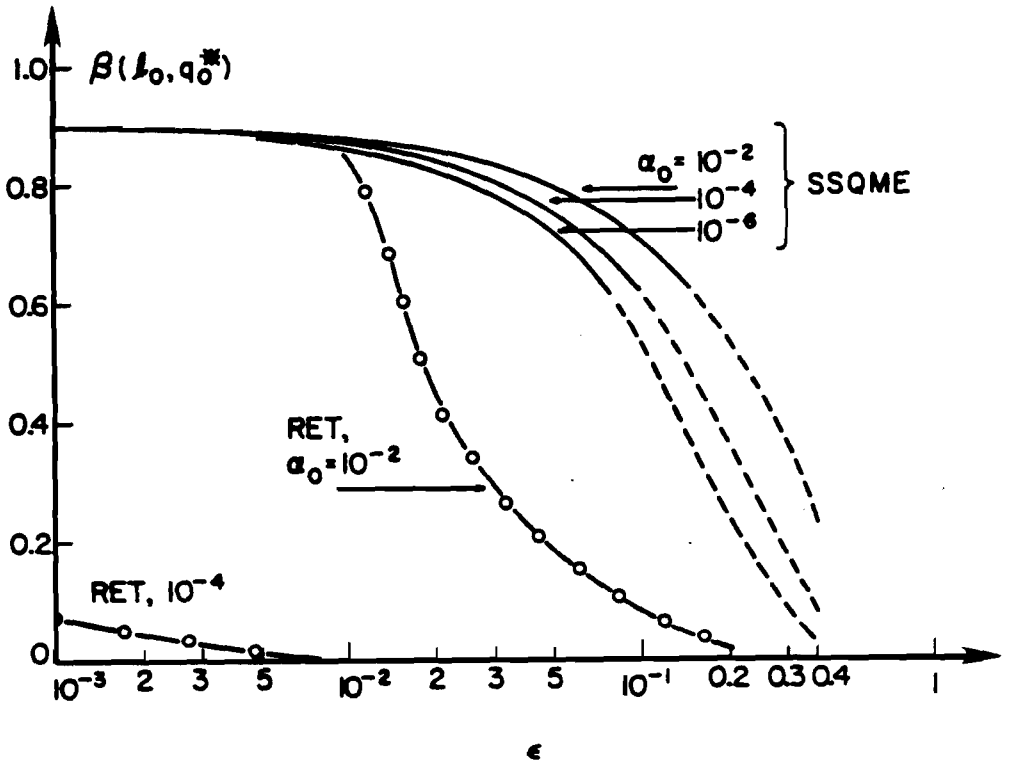
From the expression for $\beta(l_0, q_0^*)$ in Table 4.1, it is clear that the maximin bound is only slightly inferior to the detection probability in the uncontaminated case, where β is given by the same equation but with $a = A\sqrt{n}/\sigma$. This is so since for small ϵ, $rv^*(\epsilon) \triangleq V(l_0, q_0^*)/\sigma^2 \approx 1$, as can be seen in Table 4.3. (See also [3, p. 87]).

ϵ	0	.001	.002	.005	.01	.02	.05	.1	.15	.2	.25	.4
$rv^*(\epsilon)$	1	1.01	1.02	1.04	1.07	1.12	1.26	1.49	1.75	2.05	2.40	4.0

Table4.3- Upper bound on the estimation variance (and upper bound on the loss in the effective integrated SNR), $rv^*(\epsilon) \triangleq V(l_0, q_0^*)/\sigma^2$. SSQME test for narrowband signal with non-random amplitude in nearly Gaussian noise.

The corresponding maximin bound on the detection probability is depicted in Fig. 4.2 for various values of α_0, along with the results of the robust test on the envelope of section 3 (marked RET). All the curves are computed with fixed input $SNR = A^2/2\sigma^2$, such that the detection probability at the nominal Gaussian p.d.f. for all α_0's is 0.9. Very similar behavior was obtained for other SNRs. A *dramatic* improvement in performance compared to the RET is clearly observed. While the effectiveness of the RET is limited to contaminations which are smaller than roughly twice the desired false-alarm probability, (note that RET curves with $\alpha_0 < 10^{-4}$ are outside of the range of the figure) the SSQME test can protect against ϵ which is of *orders of magnitude higher* than α_0, before the detection probability decreases substantially, roughly up to $\epsilon \cong 0.1 - 0.2$. (The discontinuous portion of the curves indicate that the test may not be maximin there in accordance with Table 4.2.)

Figure 4.2- Maximin bound on the detection probability vs. ϵ, narrowband signal with non-random amplitude and unknown phase in nearly Gaussian noise. Full lines: SSQME test (discontinuous lines indicate that the test may not be maximin). Dotted lines: RET.



4.3 Random Signal Amplitude

In radar-sonar applications, modeling targets with non-random amplitude is an oversimplification. [4] Since in nominal Gaussian noise the coherent envelope test is UMP regardless of the signal amplitude statistics, it seems natural to try to also use its robust version, the SSQME test, for random amplitudes. Therefore, we will study the robustness of the SSQME test by averaging the asymptotic power of the deterministic case with respect to the assumed amplitude p.d.f. (Obviously, the false alarm probability is bounded as before.) Two cases will be treated.

a) Swerling Case 1 Target.

This is the most common radar target, composed of many independent scatterers. The amplitude is Rayleigh, $f(A) = 2A \exp(-A^2/\overline{A^2})/\overline{A^2}$, for $A \geq 0$, $\overline{A^2} \triangleq E(A^2)$. It turns out that the performance of the SSQME test is somewhat less desirable than that for the deterministic signal case.

Proposition 4.2 Under the conditions of proposition 4.1 and a Rayleigh distributed signal amplitude, the SSQME test with the nonlinearity l_0 of Eq.(4.4) satisfies $\beta(l_0,q) \leq \beta(l_0,q_0^*) \ \forall \ q \in \mathbf{P}_0$ subject to Eq.(4.8) - i.e., q_0^* is the least- favorable p.d.f. under H_0 but the *most* – favorable under H_1. However, when the nominal noise p.d.f. in $\mathbf{P}_0$ is Gaussian and l_0 is the soft limiter of Eq.(4.9), the right side inequality of Eq.(4.7) can be replaced by

$$\underset{q \in \mathbf{P}_0}{\beta(l_0, q)} \geq C \ \beta(l_0, q_0^*) \quad , \quad 0 \leq C \leq 1 \tag{4.19}$$

such that the constant C (which generally depends on ϵ, α_0 and $\beta(l_0, q_0^*)$) approaches unity as $\epsilon \to 0$, $\alpha_0 \to 0$ and as $\beta(l_0,q_0^*) \to 1$.

Proof Upon averaging with respect to A, the power of the SSQME test is easily found to be

[4] Except when the target is a perfect reflecting sphere.

$$\beta(l,\, q) = \exp\left[\frac{-t(l)/2V(l,\, q)}{1 + S(l,\, q)}\right] = \alpha_0^{1/[1\,+\,S(l,\,q)]} \tag{4.20}$$

where $S(l,\, q) \triangleq n\overline{A^2}/2V(l,\, q)$. Thus $\sup_{l\,\in\,\Psi} \beta(l,\, q_0^*) = \beta(l_0,\, q_0^*)$ since

$$\log\frac{\beta(l,\, q_0^*)}{\beta(l_0,\, q_0^*)} = \log\alpha_0\, \frac{2[V(l,\, q_0^*) - V(l_0,\, q_0^*)]}{[1 + S^{-1}(l,\, q_0^*)]\,[1 + S^{-1}(l_0,\, q_0^*)]} \leq 0 \;,\; \forall\, l \in \Psi \tag{4.21}$$

by virtue of the right side inequality in (4.12). Unfortunately, the right side inequality of (4.7) is *never* satisfied here:

$$\beta(l_0,\, q) = \alpha_0^{1/[r\,+\,S(l_0,q_0^*)]} \;,\; r \triangleq \frac{V(l_0,\, q)}{V(l_0,\, q_0^*)} \leq 1 \;\; \forall\, q \in \mathbf{P}_0 \tag{4.22}$$

Thus $\beta(l_0,\, q_0^*)$, which is given by (4.22) with $r = 1$, is always higher. This is actually the second case that was discussed in section 1.3.

However, the SSQME test will practically be *qalmost" optimal robust (in the sense that for $\forall\, q \in \mathbf{P}_0$ the performance is bounded in a narrow neighborhood: $|\,\beta(l_0,\, q_0^*) - \beta(l_0, q)\,| < \delta(\epsilon,\, \alpha_0,\, A) \to 0$, according to the discussion preceding Fig. 1.2 of section 1.3), if the right hand side of the maximin relation (4.7) could be replaced by Eq.(4.19) such that the constant C (which generally depends on ϵ, α_0 and $\beta(l_0,\, q_0^*)$) is very close to unity. From Eq.(4.22), C is obtained for the p.d.f. for which r achieves its smallest possible value over $\mathbf{P}_0$. Some algebra yields

$$\log\, C \;=\; \left[\frac{1}{1 - r_{\min}(\epsilon)}\, \frac{\log\,\alpha_0}{\log\,\beta(l_0,\, q_0^*)} - 1\right]^{-1} \log\,\beta(l_0,\, q_0^*) \tag{4.23}$$

It is clear from this expression that $c \to 1$ monotonically as $\alpha_0 \to 0$ and as $r_{\min} \to 1$. Some basic calculus shows that this convergence is also true when $\beta(l_0, q_0^*) \to 1$. At this stage we restrict the class $\mathbf{P}_0$ to the nominal Gaussian case. Using Eqs. (4.9) and (4.11)

$$\underset{q\,=\,(1\,-\,\epsilon)g_0\,+\,\epsilon h}{V(l_0,\, q)} \;=\; \frac{K^2 P_q\{\,|\,x\,|\,\geq K\} + \int_{-K}^{K} x^2 q(x)\,dx}{\left[\int_{-K}^{K} q(x)\,dx\right]^2} \tag{4.24}$$

where g_0 is the normal p.d.f. and $K(\epsilon)$ is given in (4.10). It is not difficult to see that (4.24) and hence also r are uniquely minimized when h is a point mass at the origin, as each of the terms in the numerator achieves its minimal value and the denominator its maximal value; using Eq.(4.24) with (4.10) yields

$$r_{\min}(\epsilon) = \inf_{q = (1-\epsilon)g_0 + \epsilon h} \frac{V(l_0, q)}{V(l_0, q_0^*)} = \frac{1 - \epsilon K^2 V(l_0, q_0^*)}{[1 + \epsilon V(l_0, q_0^*)]^2} \tag{4.25}$$

hence $r_{\min} \to 1$ as $\epsilon \to 0$.

	ϵ	.001	.01	.1	.2	.3	.5
	$K(\epsilon)$	2.63	1.945	1.14	.862	.685	.436
α_0	$r_{\min}(\epsilon)$	.9910	.9396	.6108	.349	.177	.028
10^{-2}	0.2	.995	.97	.78	.62	.52	.44
	$\beta^* =$ 0.5	.999	.994	.96	.93	.91	.89
	0.9	$1-2\cdot10^{-5}$	$1-1\cdot10^{-4}$	.999	.998	.998	.998
10^{-4}	0.2	.997	.98	.89	.81	.76	.72
	$\beta^* =$ 0.5	$1-5\cdot10^{-4}$	.997	.98	.97	.96	.95
	0.9	$1-1\cdot10^{-5}$	$1-7\cdot10^{-5}$	$1-5\cdot10^{-4}$	.999	.999	.999
10^{-6}	0.2	.998	.99	.93	.88	.84	.81
	$\beta^* =$ 0.5	$1-3\cdot10^{-4}$	.998	.986	.98	.97	.96
	0.9	$1-7\cdot10^{-6}$	$1-5\cdot10^{-5}$	$1-3\cdot10^{-4}$	$1-5\cdot10^{-4}$	$1-7\cdot10^{-4}$	.999
10^{-8}	0.2	.9987	.991	.95	.91	.88	.86
	$\beta^* =$ 0.5	$1-2\cdot10^{-4}$	.998	.99	.98	.98	.97
	0.9	$1-5\cdot10^{-6}$	$1-3\cdot10^{-5}$	$1-2\cdot10^{-4}$	$1-4\cdot10^{-4}$	$1-5\cdot10^{-4}$	$1-6\cdot10^{-4}$

Table 4.4 - The constant C such that $\beta(l_0, q) \geq C\ \beta(l_0, q_0^*)$, $\forall\ q \in \mathbf{P}_0$. SSQME test for Sw.1 target in nearly Gaussian noise.

The computation of (4.23) with (4.25) is summarized in Table 4.4. We observe that for the typical small α_0's and large β_0's desired in radar-sonar systems, C is almost indistinguishable from 1, even for very large contamination ϵ. Thus, for any $q \in \mathbf{P}_0$ and for the the range of practical interest, the asymptotic power is almost equal to the maximin bound, and the SSQME test is "almost robust" as conjectured. As in the deterministic amplitude case, the performance of the SSQME test is characterized by the identical decrease in the effective SNR $S(l_0, q_0^*)/(n\overline{A^2}/2\sigma^2)$, which is very small for small ϵ. Thus the qualitative behavior as in Fig. 4.2 is repeated here, with *dramatic* (asymptotic) improvement over the test of section 3.

Further thought reveals that this "almost robustness" is a consequence of always using the same fixed threshold. If it were also possible to precisely estimate the variance of the test statistic under H_0 and to adjust the threshold accordingly: $t(q^*) = -2\ln\alpha_0 V(q_0^*) \rightarrow t(q) = -2\ln\alpha_0 \hat{V}(q) \cong -2\ln\alpha_0 V(q)$, then we would obtain $\beta(q) \cong \alpha_0^{[1+S(l_0,q)]^{-1}} \geq \beta(q^*)\ \forall q \in \mathbf{P}_0$. Since in practical applications almost always the scale (variance) of the nominal noise p.d.f. is also unknown, adaptive thresholding will be unavoidable in conjunction with the robust structure to maintain CFAR. This will be treated in more detail in Section 5.2

Comparing the known lowpass signal [6] and the non-random amplitude cases with the Rayleigh amplitude case, some thought suggests that the power fails to satisfy the desired maximin relation exactly, due to an absence of a mode in the p.d.f. of the test statistic $T(l_0)$ in the present case. To verify this, we also consider the following case which does possess a mode, and represents some physical targets.

b) Swerling Case 3 Target.

This model corresponds to one dominant reflector plus Rayleigh; the amplitude p.d.f. is $f(A) = 8A^3\exp(-2A^2/\overline{A^2})/\overline{A^2}^2$ for $A \geq 0$ [17]. In this case, the desired behavior of the deterministic amplitude case is possible.

Proposition 4.3 Proposition 4.1 is valid for a Swerling 3 target model.

Proof The power of the SSQME test is found, in a manner similar to [13], by averaging the non-random amplitude result with respect to the above amplitude p.d.f., and is given by

$$\beta(l,\, q) = \left[1 + \frac{t(l)/2V(l,\, q)}{(1 + S(l,q)/2)\,(1 + 2/S(l,\, q)}\right] \exp\left[\frac{-t(l)/2V(l,\, q)}{1 + S(l,\, q)/2}\right]$$
$$= \left[1 - \frac{\log \alpha_0\, S(l,\, q)/2}{(1 + S(l,\, q)/2)^2}\right] \alpha_0^{1/[1+S(l,\, q)/2]} \tag{4.26}$$

from which $\sup_{l \in \Psi} \beta(l,\, q_0^*) = \beta(l_0,\, q_0^*)$ follows by virtue of 4.12 with some algebra. Also,

$$\beta(l_0,\, q) = \left[1 - \frac{\log \alpha_0\, S(l_0,\, q_0^*)/2}{(r\, + S(l_0,\, q_0^*)/2)^2}\right] \alpha_0^{1/[r+S(l_0,\, q_0^*)/2]} \tag{4.27}$$

where $r \triangleq V(l_0,\, q)/V(l_0,\, q_0^*) \leq 1$ as before. A sufficient condition for $\beta(l_0,\, q) \geq \beta(l_0,\, q_0^*)\ \forall r \leq 1$ is obtained from solving $d\beta(l_0,\, q)/dr \leq 0,\ \forall r \leq 1$. Upon differentiating (4.27) a quadratic inequality in $S(l_0,q_0^*)$ is obtained, whose single proper solution is given by

$$S(l_0,\, q_0^*) \geq S_C = -log\ \alpha_0 + \sqrt{\log^2 \alpha_0 + 4} \tag{4.28}$$

It is easy to see that Eq.(4.26) is monotonic increasing when $S(l_0,\, q_0^*)$ increases, hence the maximin relation (4.7) is fully satisfied whenever $\beta(l_0,\, q_0^*) \geq \beta_C$ where β_C is obtained upon substituting S_C in (4.26)

The critical values are given in Table 4.5; they are somewhat higher than those of the deterministic amplitude case, but still in the range of practical interest. Since the one-dominant-plus-Rayleigh p.d.f. corresponds to Chi-squared p.d.f. of A^2 with 2 degrees of freedom, while the non-random amplitude can be regarded as the limit of the Chi-squared family when the number of degrees of freedom N tends to infinity, it is conjectured that whenever the amplitude squared belongs to this family but $N \neq 1$ (Ray-

leigh), there exists a β_C such that when $\beta(l_0, q_0^*) \geq \beta_C$ the SSQME test is asymptotically maximin robust, and β_C decreases when N increases. It seems, however, that each case needs a separate derivation to prove this conjecture.

α_0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-8}
S_C	5.36	9.70	14.1	18.7	23.2	27.8	36.9
β_C	0.78	0.75	0.743	0.740	0.738	0.738	0.737

Table 4.5 - Critical values of $\beta(l_0, q_0^*)$ for which the SSQME test is maximin robust, Swerling 3 target.

4.4 Relationship with Weak Signal LORD

For a given observation sequence, the M-estimates defined by Eq. (4.2) are fixed numbers; they are expected to be close to zero under H_0 and also under H_1 when A is small, due to the robust properties of the estimator. Thus, an approximation is obtained by expanding the summands of (4.2) in a Taylor series in powers of $\hat{A}$, and then solving (4.2) using only the two leading terms. This yields

$$\hat{A} \cong \frac{\sum_{i=1}^{n} l(x_i)}{\sum_{i=1}^{n} l'(x_i)} \cong \frac{1}{n} \sum_{i=1}^{n} l(x_i) \tag{4.29}$$

The second approximation is justified as follows: For nominal Gaussian p.d.f. with the optimal robust nonlinearity l_0, the denominator just equals the number of x_i which are in $[-K(\epsilon), K(\epsilon)]$; it is quite close to n when ϵ is small (since $K(\epsilon)$ is then larger than the variance of the samples), and only weakly changes for different realizations of $\mathbf{x}$.

All of the locally-optimal robust detectors that have been studied in [4], [5], [9], [10] and [12] are based on this test statistic, and thus they all hide an "approximate robust estimator of zero." It is worthwhile to note that the weak result of [4] and [5], namely,

the LORD satisfies the maximin relation Eq. (1.10) on the slope of the power function at $A = 0$ only for $\alpha \geq \alpha_{\min}(\epsilon)$ is actually a consequence of this approximation. (This lower bound should be noted is *much larger than false-alarm levels of any practical interest*). Though the incremental $SNR = \lim_{A \to 0} \frac{E_1^2(\hat{A})}{VAR(\hat{A})}$ is identical with (4.16a), the numerator and denominator are somewhat different from those stated in Lemma 1

$$E_1(\hat{A}) \mid_{A \to 0} = A \; E_q(l' \,) + 0(A^2) \; ; \;\; VAR(\hat{A}) \mid_{A = 0} = \frac{1}{n} E_q(l^2) \qquad (4.30)$$

Thus there are actually two related optimization problems (for α and β), which are not parametrized by the same functional $V(l,q)$ as in the SSQME test; this lead to a different condition that must be satisfied for Eq.(1.10) to hold - see [4], [5].

The same approach could be taken for our problem. The structure of Eqs. (4.3)-(4.6) is preserved, but the M-estimators are replaced with $\hat{A}$ of (4.29). The locally-optimal robustness can be verified using similar techniques as those in [4]. Under the regularity conditions stated in [4], which require more than the conditions in (4.1), the central limit theorem is satisfied and the detection probability is again Marcum's Q-function where the parameters $\{a\,,\,b\}$ are modified according to (4.30). Using the expression for $\frac{\partial}{\partial a} Q(a\,,\,b)$ in Eq. (4.17), along with $I_1(0) = 0$, $I_1{}'(0) \neq 0$, it can be shown that the right side inequality of Eq. (1.10) is satisfied, and thus the detector is LO robust, for a limited range in the false alarm probability given by

$$\alpha_0 \geq \alpha_{\min}(\epsilon) = \min_{q \in \mathbf{P}_0} [d(q)e(q)]^{-2/|e(q)-1|} \qquad (4.31)$$

where

$$d(q) \triangleq \frac{E_q(l_0{}')}{E_{q_0^*}(l_0{}')} \;\; , \;\; d(q) \geq 1 \;\; \forall\, q \in \mathbf{P}_0 \qquad (4.32)$$

and

$$e(q) \triangleq \frac{E_{q_0^*}(l_0^2)}{E_q(l_0^2)} \;\; , \;\; e(q) \geq 1 \;\; \forall\, q \in \mathbf{P}_0 \qquad (4.33)$$

where $l_0 = -(ln\ q_0^*)'$ is the locally-optimal nonlinearity. The inequality in (4.32) is valid since l_0' is zero over the interval where the least-favorable p.d.f. q_0^* of Eq. (4.13) puts all of its contamination; the one in (4.33) is valid due to the last mentioned fact on q_0^* and to the additional regularity requirement that l_0 is monotonic nondecreasing. Because of this, the false alarm bound of Eq. (1.3) is obtained also here for $\forall\, \alpha \in (0, 1)$, but not by the same derivations as for the SSQME test. Utilizing these inequalities, it is easily verified that $\alpha_{\min}(\epsilon) \in (0, 1)$. However, we have not solved (4.31) to see if the actual values are small enough to be acceptable, as we believe that the SSQME test is a preferred solution; it is neither restricted by any $\alpha_{\min}(\epsilon)$ nor limited to weak signals. This conclusion has been verified by simulation results, to be presented in Section 5.4

5. ROBUST DETECTORS FOR REALISTIC ENVIRONMENT

The derivation of maximin robustness of the SSQME test in section 4 was based on several major assumptions, all of which have been required for previous studies of robust detection [4]-[12]. The following were assumed as completely known: a) the percentage of contamination ϵ, b) the nominal p.d.f. in the class $\mathbf{P}_0$; specifically, the variance was assumed known for the nearly Gaussian case, c) the frequency of the signal to be detected. Such knowledge concerning the detection environment is usually absent in any realistic application where robust procedures are needed. This is particularly true under the detection environment of radar (sonar) systems where the clutter (reverberation) processes rapidly change in time and space. This section extends the utility of the SSQME test for such situations, and ends with a finite-sample Monte-Carlo analysis.

5.1 Sensitivity to Unknown ϵ

In all previous works [4]-[12] the amount of contamination ϵ was assumed known. Clearly, this might not hold for a nonstationary and nonhomogeneous detection environment. Theoretically, when any knowledge of ϵ is lacking, the mixture family is no longer convex and is not well defined to allow an optimum robust test. In practice though, it might be possible to bound $\epsilon \leq \epsilon_{\max}$ based on physical considerations. (If $\epsilon_{\max}=1$, the robust estimators of the SSQME test reduce to median estimators).

An obvious but apparently unstated observation is that any density of the form $f(x;\epsilon)$ of Eq. (2.3a) belongs to the mixture family $\mathbf{P}_0(\epsilon_{\max})$ *if f* $\epsilon \leq \epsilon_{\max}$:

$$(1-\epsilon)f_0+\epsilon h = (1-\epsilon_{\max})f_0+\epsilon_{\max}[(1-\frac{\epsilon}{\epsilon_{\max}})f_0+\frac{\epsilon}{\epsilon_{\max}}h] = \cdots \tag{5.1}$$

$$\cdots \triangleq (1-\epsilon_{\max})f_0+\epsilon_{\max}h' \quad , (f_0,h,h') \in \mathbf{M}$$

and the expression in square brackets h' is a legitimate density because it is positive $(1-\epsilon/\epsilon_{\max} \geq 0)$ and integrates to 1.

Table 5.1 shows the asymptotic variances of the M-estimators of Eq.(4.2) when the input noise density is a Gauss-Gauss mixture and c^2 is the variance of the

contaminating density. They were computed from Eq.(4.11) with the soft- limiter of (4.9-4.10). In each sub table the limiter break-point k is designed according to ϵ_{max}, and the variances are given for $\epsilon \leq \epsilon_{max}$. Also indicated in the headings are the upper bounds on the estimation variance computed with the least-favorable density. It is seen (and can be shown analytically using the properties of the Gaussian c.d.f. when solving (4.11) in conjunction with (4.14)) that the variance monotonically increases with ϵ and c to the minimax bound. A similar behavior occurs with the detection probability since it is a monotone function of the effective SNR Eq.(4.16a), which is inversely proportional to the estimation variance.

The main conclusion from the table is that even when the test is designed for large ϵ_{max}, the variance is much smaller when actually $\epsilon << \epsilon_{max}$. An even more important conclusion is drawn from Table 5.2, which shows a very pessimistic case: the test is designed with $k(\epsilon_{max} = 0.5)$, and the entries in the table are the ratio of variances between those obtained with this design, and those that could be obtained if ϵ was actually known and $k(\epsilon)$ was used- $Var(k(\epsilon_{max}),\epsilon)/Var(k(\epsilon),\epsilon)$. It is clearly seen that even for ϵ=0.001 the efficiency loss is merely about 28%. This is translated to an increase of only 1.07 dB in the SNR required to achieve the same P_d. When ϵ_{max} can be more tightly bounded, the difference is even smaller.

Similar behavior is found when the contaminating density is the longer tailed Laplace density, as can be seen by dividing the entries of Table 5.3a by the corresponding ones in 5.3b.(Here however, the variance is not monotone with ϵ for small c).

To summarize this discussion, we have: A design with ϵ_{max} preserves the maximin robustness properties for $\forall \epsilon \leq \epsilon_{max}$, and the loss incurred from not knowing the actual ϵ is shown numerically to be reasonably small for the Gauss-Gauss and Gauss-Laplace mixture families considered.

Table 5.1 Estimation variances with fixed limiter $k(\epsilon_{max})$ for $\epsilon \leq \epsilon_{max}$. **Gauss–Gauss** mixture.

EPSMAX=0.005 K(EPSMAX)= 2.160 VMAX(K)=1.037

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0010	1.0066	1.0097	1.0116	1.0123	1.0125
0.0020	1.0066	1.0127	1.0167	1.0180	1.0184
0.0030	1.0066	1.0158	1.0218	1.0237	1.0243
0.0040	1.0066	1.0189	1.0269	1.0294	1.0303
0.0050	1.0066	1.0220	1.0320	1.0351	1.0362

EPSMAX=0.010 K(EPSMAX)= 1.945 VMAX(K)=1.065

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0020	1.0121	1.0177	1.0210	1.0220	1.0224
0.0040	1.0121	1.0233	1.0300	1.0321	1.0328
0.0060	1.0121	1.0289	1.0391	1.0422	1.0432
0.0080	1.0121	1.0346	1.0482	1.0523	1.0538
0.0100	1.0121	1.0403	1.0573	1.0626	1.0644

EPSMAX=0.050 K(EPSMAX)= 1.399 VMAX(K)=1.256

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0100	1.0467	1.0695	1.0807	1.0840	1.0852
0.0200	1.0467	1.0928	1.1157	1.1226	1.1250
0.0300	1.0467	1.1165	1.1518	1.1625	1.1662
0.0400	1.0467	1.1408	1.1890	1.2037	1.2088
0.0500	1.0467	1.1655	1.2274	1.2464	1.2530

Table 5.1 (continued).

Table 5.1 (continued).

EPSMAX=0.100 K(EPSMAX)= 1.140 VMAX(K)=1.490

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0200	1.0812	1.1239	1.1433	1.1490	1.1509
0.0400	1.0812	1.1685	1.2092	1.2214	1.2254
0.0600	1.0812	1.2149	1.2793	1.2989	1.3053
0.0800	1.0812	1.2634	1.3539	1.3817	1.3909
0.1000	1.0812	1.3139	1.4333	1.4706	1.4829

EPSMAX=0.200 K(EPSMAX)= 0.862 VMAX(K)=2.046

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0400	1.1393	1.2221	1.2574	1.2678	1.2706
0.0800	1.1393	1.3121	1.3904	1.4142	1.4208
0.1200	1.1393	1.4102	1.5411	1.5818	1.5933
0.1600	1.1393	1.5175	1.7125	1.7748	1.7928
0.2000	1.1393	1.6350	1.9083	1.9982	2.0249

EPSMAX=0.300 K(EPSMAX)= 0.685 VMAX(K)=2.822

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0600	1.1916	1.3162	1.3684	1.3838	1.3870
0.1200	1.1916	1.4580	1.5802	1.6177	1.6260
0.1800	1.1916	1.6200	1.8367	1.9058	1.9220
0.2400	1.1916	1.8061	2.1508	2.2655	2.2941
0.3000	1.1916	2.0214	2.5405	2.7219	2.7698

EPSMAX=0.400 K(EPSMAX)= 0.555 VMAX(K)=3.996

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0800	1.2395	1.4087	1.4794	1.5007	1.5061
0.1600	1.2395	1.6100	1.7855	1.8409	1.8557
0.2400	1.2395	1.8518	2.1838	2.2945	2.3256
0.3200	1.2395	2.1455	2.7140	2.9160	2.9761
0.4000	1.2395	2.5064	3.4395	3.7973	3.9099

EPSMAX=0.500 K(EPSMAX)= 0.436 VMAX(K)=5.928

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.1000	1.2916	1.5086	1.5999	1.6267	1.6294
0.2000	1.2916	1.7789	2.0195	2.0949	2.1047
0.3000	1.2916	2.1211	2.6094	2.7747	2.8021
0.4000	1.2916	2.5626	3.4734	3.8124	3.8808
0.5000	1.2916	3.1447	4.8072	5.5038	5.6726

Table 5.2 ratio of estimation variances between designs with fixed $k(\epsilon_{max}=0.5)$ and variable $k(\epsilon)$,**Gauss–Gauss** mixture.

VAR(K(EPSMAX))/VAR(K(EPS)) EPSMAX=0.500 K(E PSMAX)= 0.436

F=(1-EPS)*N(0,1)+EPS*N(0,C**2)

EPS/C	1	3	10	30	100
0.0010	1.2897	1.2868	1.2838	1.2828	1.2824
0.0050	1.2832	1.2734	1.2645	1.2617	1.2604
0.0100	1.2762	1.2605	1.2471	1.2428	1.2409
0.0500	1.2340	1.1963	1.1687	1.1600	1.1547
0.1000	1.1946	1.1482	1.1162	1.1062	1.0987
0.2000	1.1337	1.0880	1.0582	1.0484	1.0394
0.3000	1.0839	1.0493	1.0271	1.0194	1.0117
0.4000	1.0420	1.0224	1.0099	1.0040	1.0031
0.5000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5.3 Estimation variances for **Gauss–Laplace** mixture. a) Fixed $k(\epsilon_{max})$. b)Variable $k(\epsilon)$

EPSMAX=0.500 K(EPSMAX)= 0.436 VMAX(K)=5.928

F=(1-EPS)*N(0,1)+EPS*LAPLACE(0,C**2)

EPS/C	1	3	10	30	100
0.1000	1.1863	1.4343	1.5708	1.6172	1.6379
0.2000	1.0922	1.5996	1.9403	2.0676	2.1252
0.3000	1.0076	1.7924	2.4425	2.7139	2.8392
0.4000	0.9314	2.0189	3.1478	3.6848	3.9411
0.5000	0.8626	2.2874	4.1795	5.2347	5.7638

VARIANCES WITH OPTIMAL K(EPS)

F=(1-EPS)*N(0,1)+EPS*LAPLACE(0,C**2)

EPS/C	1	3	10	30	100
0.1000	1.0438	1.2455	1.3979	1.4570	1.4799
0.2000	1.0318	1.4711	1.8222	1.9647	2.0244
0.3000	0.9914	1.7138	2.3675	2.6536	2.7837
0.4000	0.9327	1.9821	3.1101	3.6591	3.9318
0.5000	0.8626	2.2874	4.1795	5.2347	5.7638

5.2 Extension of the SSQME Test for Unknown Scale of the Nominal Density

The scale σ of the nominal density in the ϵ-mixture family affects the implementation and performance of the SSQME test in two ways. First, σ must be known for construction of the optimal robust M-estimates, as the break points of the nonlinearity are functions of it, Eq.(4.14); e.g., they are equal to $\pm\sigma k(\epsilon)$ for the nominal Gaussian case. Second, the detection threshold setting $t(l_0)$ must be proportional to $V(l_0, q_0^*)$ to obtain the desired α_0 level, (see Table 4.1), and V itself is proportional to σ^2. Even if the optimal M-estimators are replaced by some other robust and scale invariant estimators of the quadrature locations, the second crucial problem must still be addressed.

In the specific case of nominal Gaussian density, which is the main interest of our work, the first problem is solved together with a substantial simplification in implementation complexity by utilizing Tukey's alpha-trimmed mean estimator instead of the M-estimator. This estimate is defined by:

$$\overline{X}_\alpha = \frac{1}{n(1-2\alpha)} \left[\sum_{i=2+\lfloor n\alpha \rfloor}^{n-1-\lfloor n\alpha \rfloor} x_{(i)} + (1+\lfloor n\alpha \rfloor - n\alpha)[x_{(1+\lfloor n\alpha \rfloor} + x_{(n-\lfloor n\alpha \rfloor)}] \right] \tag{5.2}$$

where $x_{(i)}$ are the ordered samples: $x_{(1)} \leq x_{(2)} \leq \ldots\ldots x_{(n)}$, and $[n\alpha]$ is the greatest integer in $n\alpha$ for $0 \leq \alpha < 0.5$. The estimator simply deletes $[n\alpha]$ samples from each end and then takes the weighted mean of the remaining. Its advantage over Huber's M-estimator is that solving an implicit nonlinear equation is not necessary. With the above weighting it is translation invariant for all n (and unbiased when $E_f(x_i)$ exists). Bickel [33] showed that it is asymptotically Gaussian and its variance for symmetric f is

$$Var(\sqrt{n}\,\overline{X}_\alpha) \triangleq \sigma_\alpha^2 = \frac{2}{(1-2\alpha)^2} \left[\int_0^{x_{1-\alpha}} x^2 f(x)dx + \alpha x_{1-\alpha}^2 \right] \tag{5.3}$$

where $F(x_{1-\alpha}) = 1-\alpha$. The two estimators are quite similar for proper choice of α and $k(\epsilon)$; they both sum linearly the main bulk of the samples, but Huber's estimator treat

large samples and outliers by limiting them, while the α-trimmed estimator censors them completely. The main difference is that for small n the α -trimmed will always discard some of the observations, even if they all are small and come from the nominal, while the other might not. For $n \rightarrow \infty$, however, they are equivalent in quadratic mean. The variance of Huber's estimator with the soft- limiter nonlinearity $l_0(x;-k,k)$ is

$$V(l_0,f) = \frac{2[\int_0^k x^2 f(x)dx + k^2 F(-k)]}{[1-2F(-k)]^2} \tag{5.4}$$

This is identical to (5.3) when one choses $\alpha(\epsilon)=F(-k(\epsilon))$. Hence they both have the same asymptotic distribution and minimax property so that the scale invariant α-trimmed mean can be substituted in the SSQME test without change in the derivation.

Two different asymptotically scale invariant extensions of the SSQME test are next discussed. The difference between them depends on the availability of a "noise-reference" channel.

a) "Sliding-Window" SSQME test

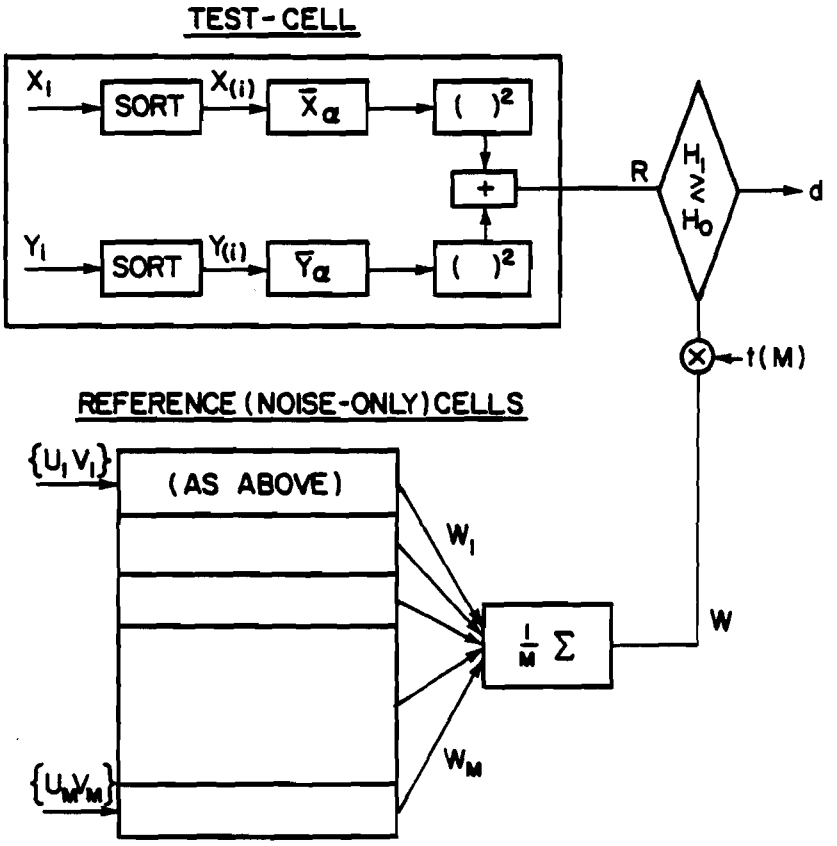
We follow the definitions and the physical setup of section 2.2. In addition to the "test-cell" samples $\{\mathbf{x},\mathbf{y}\}$, there are also available M "noise- reference" vectors of observations: $\{u_{ij},v_{ij}\}, i = 1,...n, j = 1,...,M$. $\mathbf{u}_j \triangleq \{u_{ij}\}_{i=1}^n, j = 1,...M$ are statistically independent vectors that are identically distributed as $\mathbf{x}$ is under H_0, and likewise for $\mathbf{v}_j$ and $\mathbf{y}$. An adaptive threshold test is constructed by using the SSQME structure in both the "test-cell" statistic $R(\mathbf{x},\mathbf{y})$ and in the threshold estimator $W(\mathbf{u},\mathbf{v})$ of Eqs.(2.8)-(2.11). Specifically, the test is

$$DECIDE \quad H_1 \quad if f \quad R_\alpha(\mathbf{x},\mathbf{y}) \quad \geq \quad t(M)W_\alpha(\mathbf{u},\mathbf{v}) \tag{5.5}$$

where $R_\alpha(\mathbf{x},\mathbf{y})$ is identical with $T_n(l_0)$ of Eq.(4.3) with the M-estimators $\hat{A}_x$ and $\hat{A}_y$ replaced by the α– trimmed means of the $\mathbf{x}$ and $\mathbf{y}$ samples, respectively. In a similar manner, $W_\alpha(\mathbf{u},\mathbf{v})=(1/M)\sum_{j=1}^{M} W_{j\alpha}$, where $W_{j\alpha}=\overline{U}_{j\alpha}^2+\overline{V}_{j\alpha}^2$; $\overline{U}_{j\alpha}$ and $\overline{V}_{j\alpha}$ are the α-

trimmed means of $\mathbf{u}_j$ and $\mathbf{v}_j$, respectively. The test structure is shown in Fig. 5.1.

Fig. 5.1 Sliding-Window scale invariant SSQME robust test



Under H_0, for $n \rightarrow \infty$ the $W_j \; \forall j$ and R are not only identically distributed (this is true also for finite n), but the unknown input scale appears only through $V(l_0,q)$, which is again a scale parameter of the (one sided exponential) distributions of these statistics. Hence, the test (5.5) is scale invariant and CFAR: $\alpha_0 = Prob\{R \geq tW\} \; \forall$ $q \in \mathbf{P}_0$. It is important to emphasize that here, unlike the fixed threshold SSQME test, the multiplier $t(M)$ *is not* a function of $V(l_0,q_0^*)$, so that the same α_0 is obtained for $\forall$ $q \in \mathbf{P}_0$. Moreover, the expected value of the statistic W_j is exactly equal to the variance $V(l_0,q)$ of the estimators in the "test-cell" by definition, and, in addition, when $M \rightarrow \infty$ the estimation of this variance by the "sliding-window" sample mean $W(\mathbf{u},\mathbf{v})$ converges almost surely to the true value. Hence, the test (5.5) will be asymptotically $(M \rightarrow \infty)$ a.s. equivalent to the SSQME test when σ^2 of the nominal density is known. For finite M the false alarm probability is given by $\alpha_0 = (1+t(M)/M)^{-M}$, and the detection probability of a Rayleigh amplitude signal by (compare [18]-[20] for the pure Gaussian case)

$$\beta(l_0,q) = [1+t(M)/M(1+S(l_0,q)]^{-M} \tag{5.6}$$

Here $S(l_0,q) = nA_0^2/2\sigma^2 V(l_0,q)$ is the effective SNR computed at the optimal non-linearity (or its corresponding trimming ratio α) for density q [1]. Note that the limit of these expressions with $M \rightarrow \infty$ are indeed the exponential probabilities which are valid for the SSQME test as in section 4.3 case a), and that for $M \geq 16$ the detectability loss incurred by the adaptive threshold scheme is very small, c.f. [18]-[20].

The monotonity of $\beta(l_0,q)$ in $S(l_0,q)$ ensures as before that the upper bound on the variance in $q \in \mathbf{P}_0$ is translated to a lower bound on the detection probability. Note that the problem with the Rayleigh signal of section 4.3 case a) is solved here, as $t(M)$ is not a function of $V(l_0,q_0^*)$ as it was there. Hence $\beta(l_0,q) \geq \beta(l_0,q_0^*) \; \forall \; E(A^2) > 0$! Moreover, this conclusion remains valid for the various amplitude models discussed in

[1] The expression for deterministic signal amplitude can be derived as in [18] with the effective SNR as given by (4.16a)

section 4, and in general, for any amplitude p.d.f. that leads to a monotone increasing $\beta(S \mid W)$ in S for the known scale case. (Since $\beta(S)=E_W \beta(S \mid W)$ is obviously monotonic).

We can summarize this discussion in the following proposition:

Proposition 5.1 Let $\mathbf{D}_M$ be the class of all hypothesis tests of the form $R(T_x, T_y) \geq t_M W(T_{u_j}, T_{v_j}, j=1, \cdots M)$ where W and R are as in Eq. (5.5), with the α-trimmed estimators replaced by any other translation invariant statistics T, and let $d(\alpha) \in \mathbf{D}_M$ be the test with α-trimmed estimators. Let $\mathbf{D}_\infty$ be the class of all decision rules based on $\{T_x, T_y\}$ - i.e., without imposing the structure of Eq. (5.5) and precluding reference samples. Then --

a) The test is CFAR for any $q \in \mathbf{P}_{0\sigma}$ and $P_{fa}=(1+t(M)/M)^{-M}$.

b) If $\alpha^*(\epsilon)=F_0^*(-k(\epsilon))$, where F_0^* is the c.d.f. of q_0^*, q_0^* and $d^*(\alpha^*(\epsilon))$ are a saddle-point pair in $(\mathbf{P}_{0\sigma}, \mathbf{D}_M)$ for the detection probability

$$\underset{d \in \mathbf{D}_M}{\beta}(q_0^*, d) \quad \leq \quad \beta(q_0^*, d^*) \quad \leq \quad \underset{q \in \mathbf{P}_{0\sigma}}{\beta}(q, d^*) \tag{5.7}$$

i.e., the test is asymptotically maximin robust. Moreover, as $\epsilon \to 0$, $\beta(q_0^*, d^*) \to$ monotonely to the detection probability of the best scale invariant test against pure Gaussian noise. Thus, small deviations from the assumed nominal model only slightly degrade the performance. $\beta(q_0^*, d^*)$ is given by Eq. (5.6) evaluated at q_0^*.

c) When also $M \to \infty, d^*$ is equivalent to the Neyman-Pearson test in the wider class $\mathbf{D}_\infty$ for $q_0^*(\sigma=1)$ - i.e., assuming that σ is known and only $\{T_x, T_y\}$ are available as observables.

For the general case of a non- Gaussian *nominal* density f_0, the α– trimmed mean is not the optimal estimator, although it is still robust in Hampel's sense as will be discussed in the following. All the previous conclusions remain valid, if the α -trimmed estimators are replaced by the optimal L-estimators for the same least favorable p.d.f. of Eq. (4.13). The location invariant and scale equivariant L-estimators are given by

$L_n(\mathbf{x}) = \sum a_{ni}\, x_{(i)}$, etc., where the weights are generated from a score function m(s) according to

$$a_{ni} = \int_{(i-1)/n}^{i/n} m(s)\,ds \tag{5.8}$$

and $m(\cdot)$ is related to q_0^* of Eq.(4.13) through

$$m(F_0^*(x)) = \begin{cases} \dfrac{1}{I(F_0^*)} \dfrac{d^2}{dx^2}(-\log f_0(x;1)) & , \ |x| \le x_1 \\ 0 & , \ otherwise \end{cases} \tag{5.9}$$

where

$f_0(x_1;1)/k - F_0(-x_1;1) = \epsilon/2(1-\epsilon)$ and $|f_0'(x_1;1)/f_0(x_1;1)| = k$. Notice that also here the censoring percentage is $F_0^*(-x_1)$. The minimax relation on the estimator's variance remain valid (where $V(q_0^*) = 1/I(q_0^*)$), due to the asymptotic equivalence between L and M-estimators [Jaeckel, 46], provided the a_{ni} are chosen as above. (Stigler [47] proved later that if the nominal p.d.f is any symmetric, twice differentiable and strongly unimodal p.d.f. on R^1, the L-estimator corresponding the above score function is indeed asymptotically Gaussian with the variance assumed in [46]).

To summarize, the SW adaptive threshold extension of the SSQME test is a suitable choice for radar-sonar systems where "noise-reference" samples are conveniently available. It possesses all the asymptotic desired maximin properties; its implementation in the nominal Gaussian case is relatively easy based on α-trimmed estimators which do not require more than ordering of the data, an operation that is common to most non-parametric tests and is becoming available for real time implementation with VLSI technology.

b)SSQME test with a preliminary scale estimate

We assume now that $\{\mathbf{x},\mathbf{y}\}$ are the only available observables, and the nuisance scale estimate must be derived from the "signal plus noise " data itself. We draw and build on Huber's theory [3] of simultaneous estimation of location and scale. For

completeness of the presentation, we next quote and make some interpretations on Huber's main results:

i) Simultaneous M-estimates for location and scale are a pair of statistics (T_n, S_n) determined by solving equations of the form

$$\sum_{i=1}^{n} \Psi\left(\frac{x_i - T_n}{S_n}\right) = 0 \tag{5.10a}$$

$$\sum_{i=1}^{n} \chi\left(\frac{x_i - T_n}{S_n}\right) = 0 \tag{5.10b}$$

These are similar to the ML estimates of θ and σ for a family of densities $\frac{1}{\sigma} f\left(\frac{x-\theta}{\sigma}\right)$, with proper selection of Ψ and χ. As most common test statistics and estimators depend on the samples only through the empirical distribution function $F_n(\mathbf{x}\,;\,x) \triangleq \frac{1}{n}\sum I_{\{x_i < x\}}$, where $I_{\{A\}}$ is the indicator function of the set A, it is convenient to express $T_n(\mathbf{x}) = T(F_n)$ and $S_n(\mathbf{x}) = S(F_n)$ in terms of the functionals defined by:

$$\int \Psi\left(\frac{x - T(F)}{S(F)}\right) F(dx) = 0 \tag{5.11a}$$

$$\int \chi\left(\frac{x - T(F)}{S(F)}\right) F(dx) = 0 \tag{5.11b}$$

Hampel's [34] influence curve (IC) is a very useful heuristic tool of robust statistics, which describes the (suitably normed) limiting influence of a single observation on the estimator. Intuitively, a qualitative robust estimator *must have a bounded IC.* If $\sqrt{n}\,[T(F_n) - T(F)]$ is asymptotically zero mean Gaussian, its variance is given by (using (5.13) below)-

$$V(F,T) = \int IC^2(x;F,T)F(dx) = S^2(F)\frac{\int \Psi^2\left(\frac{x}{S(F)}\right)F(dx)}{\left[\int \Psi'\left(\frac{x}{S(F)}\right)F(dx)\right]^2} \tag{5.12}$$

with a similar expression for $V(F,S)$. If (and only if) F and χ are symmetric, and Ψ is skew-symmetric, the IC's of the simultaneous estimates are uncoupled; in par-

ticular, only S(F) but neither the IC nor the asymptotic variance of S enters into the expression of IC(T). The IC's are given by:

$$IC(x;F,T) = \frac{\Psi(\frac{x}{S(F)})S(F)}{\int \Psi'(\frac{x}{S(F)})F(dx)} \tag{5.13}$$

$$IC(x;F,S) = \frac{\chi(\frac{x}{S(F)})S(F)}{\int \chi'(\frac{x}{S(F)})\frac{x}{S(F)}F(dx)} \tag{5.14}$$

Note that for a given distribution, T(F) and S(F) are numbers; T_n and S_n are Fisher consistent estimates if T(F) and S(F) are equal to the estimated parameters. Under the above mentioned symmetry conditions, the estimates are asymptotically uncorrelated and hence independent.

ii) The asymptotic robustness (in Hampel's sense) is clear from (5.13)-(5.14), provided that Ψ and χ are bounded functions. In section 6.4 of [3] it is shown that under relatively mild conditions the coupled estimates are consistent $(T_n,S_n) \xrightarrow{\mathbf{P}} (T(F),S(F))$ and jointly Gaussian.

iii) The IC of T for the case when the nominal scale is known is given by the same expression as (5.13) but with S(F)=1. If S(F) does not change too much when $F \in \mathbf{P}_0$ (a property which must be satisfied for any reasonably consistent and robust estimate of the scale), the desired properties of the location estimate with known scale will be roughly preserved. However, it is not possible to obtain an exact minimax, since the loss of the translation invariance symmetry in the multi-parameter case does not enable extension of the parametrization of the nominal model throughout a convex uncertainty neighborhood, c.f. [3, section 11.1]. It should also be clear that the variational techniques of [3] are not applicable for finding the density that minimizes $V(F,T)$, even when the asymptotic variance of S is considered as a nuisance factor, due to the deep and implicit nonlinear coupling

of Eq.(5.12) subject to Eqs.(5.11).

iv) The simultaneous solution of equations (5.10) is perhaps an unnecessary complicated. Simplified variants are the *one step M-estimates*. They are obtained by starting with some preliminary estimates $T_n^{(0)}$ and $S_n^{(0)}$, and then solving Eq.(5.11a) approximately by applying Newton's method *just once*. The following variant emerged in the extensive "Princeton Monte-Carlo study " of Andrews et.al. [35] as very robust, even for small sample sizes. The preliminary estimates are the median: $T_n^{(0)}(\mathbf{x}) = med\{x_i\}$, and the median of the absolute deviations from it (MAD): $S_n^{(0)}(\mathbf{x}) = med\{\mid x_i - T_n^{(0)}(\mathbf{x}) \mid\}/0.6745$, where $0.6745 \cong \Phi^{-1}(3/4)$ to get consistency at the Gaussian ($\chi(x) = sign(\mid x \mid -1)$, thus $S^{(0)}(F) = F^{-1}(3/4)$ for a symmetric density from (5.11b)). The MAD is the limiting case ($\epsilon \rightarrow 1$) of the minimax robust scale estimate, and its computation is simpler. Then,

$$T_n^{(1)}(\mathbf{x}) = T_n^{(0)}(\mathbf{x}) + \frac{\frac{1}{n}\sum \Psi(\frac{x_i - T_n^{(0)}}{S_n^{(0)}}) S_n^{(0)}}{\frac{1}{n}\sum \Psi'(\frac{x_i - T_n^{(0)}}{S_n^{(0)}})} \tag{5.15}$$

$T_n^{(1)}$ is asymptotically ($n \rightarrow \infty$) equivalent to the full solution of Eq.(5.10a) $T_n^{(\infty)}$, provided the previously mentioned symmetry conditions are satisfied. Moreover, with these symmetry conditions, any one-step estimator with translation invariant and odd $T_n^{(0)}$: $T_n^{(0)}(\mathbf{x}+c) = T_n^{(0)}(\mathbf{x}) + c$, $T_n^{(0)}(-\mathbf{x}) = -T_n^{(0)}(\mathbf{x})$), will have the same IC (and hence identical asymptotic properties) as the full solution of the coupled equations. Note that for the nominal Gaussian case the computation of (5.15) is of the same order of complexity as of the nonlinear detectors typical of the LORD approach (section 1.3), provided $T_n^{(0)}$ and $S_n^{(0)}$ are available after two orderings of the data (the denominator is just the number of normalized samples in $[-k, k]$).

With this background , it seems natural to preserve the structure of the SSQME test, but with scale invariant location estimators and an adaptive threshold which is

derived from the scale estimate. Thus we propose the following test:

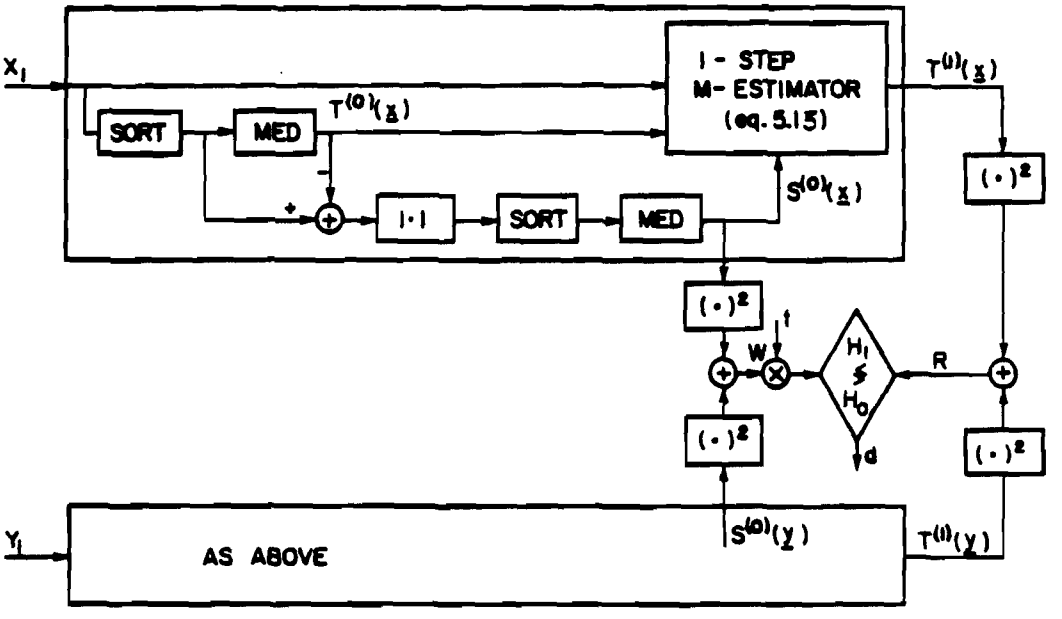
$$DECIDE \quad H_1 \quad iff \quad R(\mathbf{x},\mathbf{y}) \geq t_n W(\mathbf{x},\mathbf{y}) \tag{5.16}$$

where

$$R(\mathbf{x},\mathbf{y}) \triangleq [T_n^{(1)}(\mathbf{x})]^2+[T_n^{(1)}(\mathbf{y})]^2 \ , \ W(\mathbf{x},\mathbf{y}) \triangleq [S_n^{(0)}(\mathbf{x})]^2+[S_n^{(0)}(\mathbf{y})]^2 \tag{5.17}$$

$T_n^{(1)}$ is the one step M-estimator of Eq.(5.15) and $S_n^{(0)}$ is the MAD scale estimator. The structure of the test is shown in Fig. 5.2.

Fig. 5.2 SSQME test with MAD preliminary scale estimate.



The asymptotic properties of the test are summarized in the following. They are valid for all simultaneous and one step M-estimators provided the symmetry properties that were mentioned before hold.

Proposition 5.2 a) The test (5.16) is CFAR for any fixed density $f \in \mathbf{P}_0$ with unknown (variable) σ. b) With variable f , the changes in the asymptotic false-alarm probability are governed only by the normalized variance of the location estimates : $V'(F, T^{(1)}) \triangleq V(F, T^{(1)})/(S^{(0)}(F))^2$, but not by the variance of the scale estimate. c) The asymptotic detection probability is given by the same expression as for a fixed threshold SSQME test: $\beta(F)=\alpha(F)^{1/(1+SNR(F))}$ (for a Rayleigh signal), where

$$SNR(F) = \frac{nA_0^2}{2[S^{(0)}(F)]^2 V'(F, T^{(1)})} \tag{5.18}$$

is also not a function of the scale estimator variance.

Proof 1) Both $T_n^{(1)}$ and $S_n^{(0)}$ are scale invariant: $S_n^{(0)}(\sigma \mathbf{x})=\sigma S_n^{(0)}(\mathbf{x})$, etc., from which a) follows immediately. Also, $S^{(0)}(F_\sigma) = \sigma S^{(0)}(F_{\sigma=1})$, hence we will omit the scale subscript on F for convenience.

2) The probability of acceptance of H_1 is given by $\int_0^\infty f_W(w)dw \int_{t_n w}^\infty f_R(r)dr$, since R and W are asymptotically independent when the symmetry conditions of i) above are met. R is one sided exponentially distributed, with

$$E_0(R) = \frac{2}{n}(S^{(0)}(F))^2 V'(F, T^{(1)}) \quad , \quad E_1(R) = E_0(R)[1+SNR] \tag{5.19}$$

by straightforward calculation. Note that the definition of SNR is justified (by analogy to the classical case) as $V'(F, T^{(1)})$ is also σ invariant by its definition, as long as the scale and location estimates are also invariant. Hence the input noise variance affects SNR only through $S^{(0)}(F_\sigma)=\sigma F^{(0)}(F_1)$. As W is the sum of the squares of two independent identically distributed Gaussian r.v.'s $N(a, v^2)$, it has a *noncentral* χ^2 distribution with 2 degrees of freedom and noncentrality parameter $\lambda=2a^2$, c.f. [13]. Explicitly,

$$f_W(w)=\frac{1}{2v}\exp(-\frac{w+2a^2}{2v})I_0(\frac{a\sqrt{2w}}{v}) \tag{5.20}$$

where I_0 is the modified Bessel function of zeroth order. Here $a=S^{(0)}(F)$ and $v=\frac{1}{n}V(F,S^{(0)})$ under H_0 and under H_1, as $S_n^{(0)}(\mathbf{x}+c)=S_n^{(0)}(\mathbf{x})$. The desired probabilities are thus computed from:

$$p_i \quad = \quad \frac{\exp(-a^2/v)}{2v}\int_0^\infty \exp[-w(\frac{1}{2v}+\frac{t_n}{E_i(R)})]I_0(\frac{a\sqrt{2w}}{v})dw \quad , \quad i=H_0, H_1 \tag{5.21}$$

The evaluation of this integral is facilitated by-

$$\int_0^\infty \exp(-x)I_0(2\sqrt{yx})dx \quad = \quad \exp(y) \tag{5.22}$$

which can be derived by a change of variables from the normalization of the Rician density. The result is

$$p_i \quad = \quad \frac{\exp[-\frac{2a^2t_n}{E_i(R)}(1+2t_n v/E_i(R))^{-1}]}{1+2t_n v/E_i(R)} \quad , \quad i=H_0, H_1 \tag{5.23}$$

As $a^2/E_i(R)\approx n$ and $v/E_i(R)$ is not a function of n, in order to get a non-zero α as $n\rightarrow\infty$ we must have $t_n=t/n$. Thus when $n\rightarrow\infty$ the terms that contain $V(F,S^{(0)}(F))$ through v cancel out and $\lim_{n\rightarrow\infty} p_0 = \exp(-t/V'(F,T^{(1)})$ follows and proves b). The proof of c) is similar upon substitution of $E_1(R)$ from Eq.(5.19) into (5.23) and taking the limit as $n\rightarrow\infty$.

Comments and interpretation

1)Since the error probabilities are not functions of the variance of the scale estimator, $S_n^{(0)}$ should be chosen to achieve better and flatter consistency over $\mathbf{P}_o$, and not according to its variability. 2) The false alarm probability with the preliminary estimate is parametrized by $V'(F,T^{(1)})$, which is roughly equal to the variance under the known scale case, if $S^{(0)}(F)\cong 1$ and it does not change too much over $\mathbf{P}_0$. Hence we have a similar situation as in Chapter 4, and the upper bound on the false alarm probability is approximately satisfied if the least favorable density q_0^* is chosen. While it does not

seem possible to give an analytic characterization of the density that maximizes $V'(F,T^{(1)})$ over the ϵ–mixture family $\mathbf{P}_0$ as was discussed in iii) above, if the noise density is restricted to a narrower family (e.g., Gauss-Gauss or Gauss-Laplace ϵ– mixtures), the supremum of the variance can be found (at least numerically) and adjusted accordingly to obtain an upper bound on the α over the restricted family. 3)With that restricted maximization, an approximate upper bound on the detection probability will also be obtained as follows. Denote by $\mathbf{P}_{0r}$ the restricted noise family. Since $S^{(0)}(F,D)=F^{-1}(3/4)/D$ (for the MAD estimator) changes over $\mathbf{P}_{0r}$, its arbitrary normalization D can be used as an optimization parameter. Let $V^{**}(D)=\underset{\mathbf{P}_{0r}}{Max}V'(F,T^{(1)},D)$ and $\alpha_0=\exp(-t/V^{**}(D))\geq\alpha(F,D)$. Under the same approximation of section 4.3 for sufficiently high desired probability of detection ($\beta\gtrsim 0.5$), with substitution of SNR from Eq. (5.18):

$$\beta(F,D) = \alpha(F,D)^{1/(1+SNR(F))} \cong \alpha_0^{2[S^{(0)}(F,D)]^2V^{**}(D)/nA_0^2} \tag{5.24}$$

Hence the additional detectability loss from scale estimation relative to the worst case where σ is known is characterized by the effective SNR loss

$$L(F,D) = \frac{V^{**}(D)[S^{(0)}(F,D)]^2}{V(l_0,q^*)} \tag{5.25}$$

and D^{**} can be chosen to minimize the maximum loss over $\mathbf{P}_{0r}$: $D^{**}=\underset{D}{ArgMin}\{\underset{f\in\mathbf{P}_{0r}}{Max}L(F,D)\}$. Hence, α (β) is bounded from above (below) over the family, and the additional loss due to scale estimation is low if $L(F^{**},D^{**})\approx 1$.

Figures 5.3a-5.3f show the results of such a numerical optimization over the Gauss-Gauss mixture family, vs. the r.m.s. power of the contamination c. In each figure the dashed lines are of $V'(F,T^{(1)})/V^{**}$, and the continuous lines are of $L(F,D)$. One set of graphs in each figure corresponds to the optimization with respect to D, while the other was obtained where D was taken to get consistency at the least-favorable density of the known σ case, i.e., $D^*=F^{*(-1)}(3/4)$. Using Eq.(4.10), it is given by

$$D^* = \begin{cases} k - \frac{1}{k}\ln\left[\frac{k}{4(1-\epsilon)f_0(k)}\right] & , \quad \frac{f_0(k)}{k} \geq \frac{1}{4(1-\epsilon)} \\ \Phi^{-1}\left[\frac{1+2(1-\epsilon)}{4(1-\epsilon)}\right] & , \quad \frac{f_0(k)}{k} < \frac{1}{4(1-\epsilon)} \end{cases} \tag{5.26}$$

The figures demonstrate that there is not any *additional* loss (above that incurred from $V(l_0,q_0^*)/\sigma^2$) from the scale estimation for $\epsilon \leq 0.2$, that the loss is negligible (about 4%) for $\epsilon=0.3$, and it is quite small (about 26%,corresponding to 1 dB) for $\epsilon=0.4$. Even for the huge uncertainty $\epsilon=0.5$ the loss is reasonable (50%) if $c \leq 20$. For $\epsilon>0.5$, it seems that $L(F,D^{**})$ is unbounded. This agrees with the "breakdown" point of the joint estimation, as evaluated by Huber [3, pp. 141-146]. Also, the loss difference between the optimal D^{**} and the normalization D^* is quite small. As $D^{**}>0.6745$, the efficiency at the nominal Gaussian is somewhat sacrificed to obtain better performance over $\mathbf{P}_{0r}$ (compare with Table 5.1). Similar conclusions can be drawn from Figures 5.4a-5.4d which are valid for the Gauss-Laplace mixture family. When the nominal p.d.f. is Gaussian, the test of Eq.(5.17) can be simplified by utilizing α-trimmed estimators in place of the one step location estimator. In that case, it turns out that the agreement with the performance of the known scale is even closer, see [48].

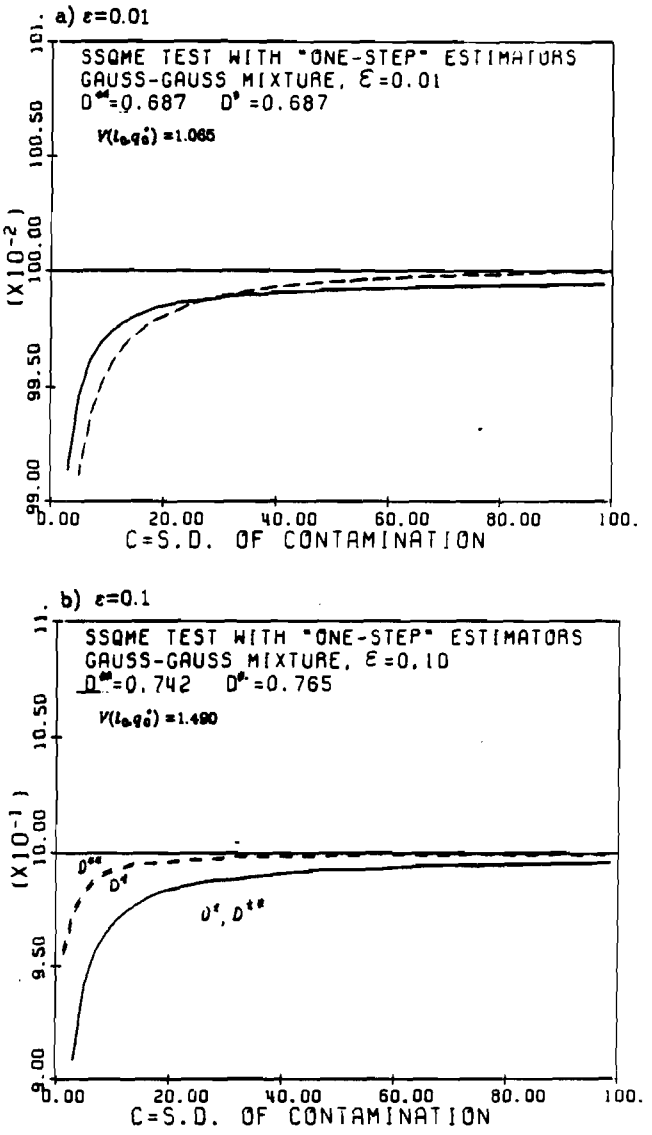
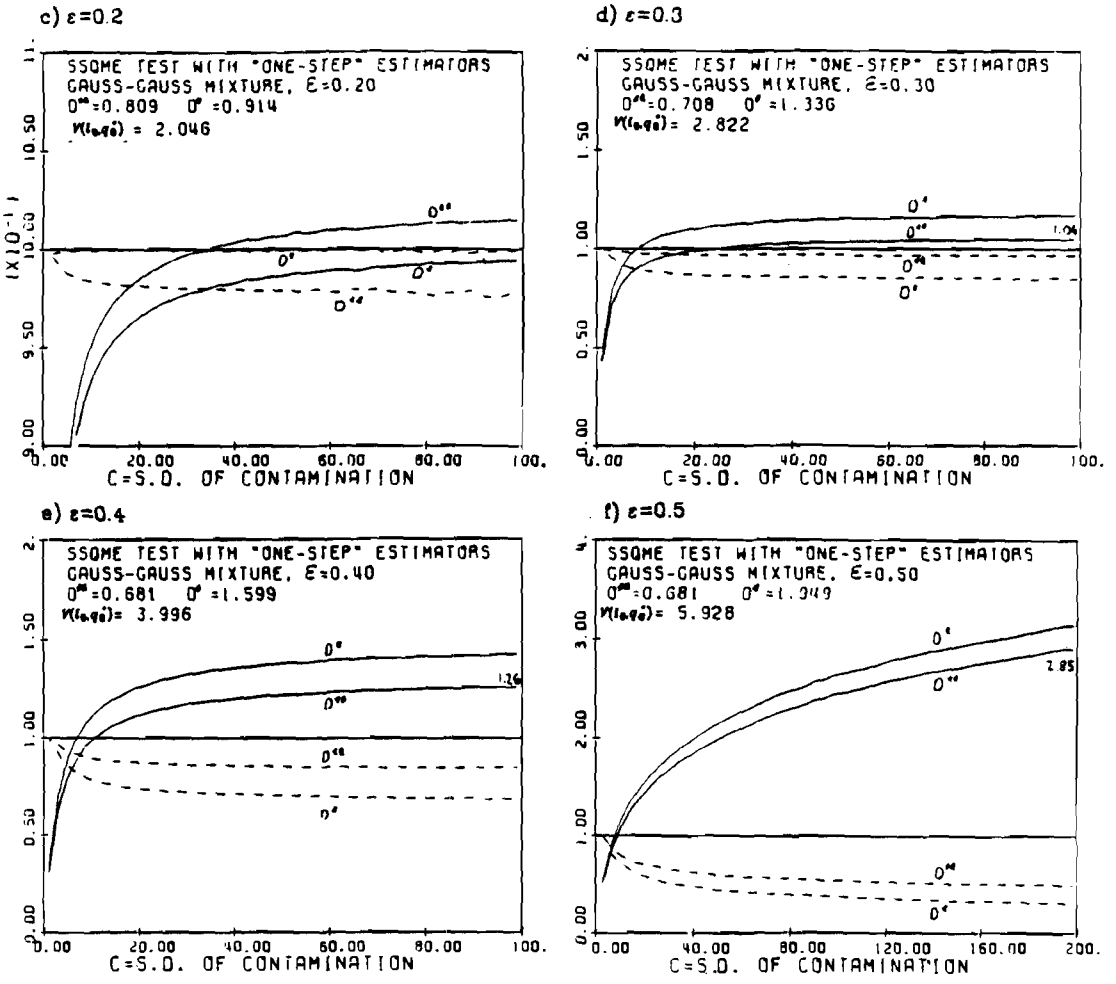


Fig. 5.3 Normalized variance $V'(F,T^{(1)})/V^{**}$ (dashed lines) and effective SNR loss $L(F,D)$ of SSQME test with a preliminary scale estimate. Gauss- Gauss mixture, c^2=contamination variance. D^* - scale normalization of the known-σ least favorable density, D^{**} - optimal scale normalization. a) ϵ=0.01 .

Fig. 5.3 (continued)



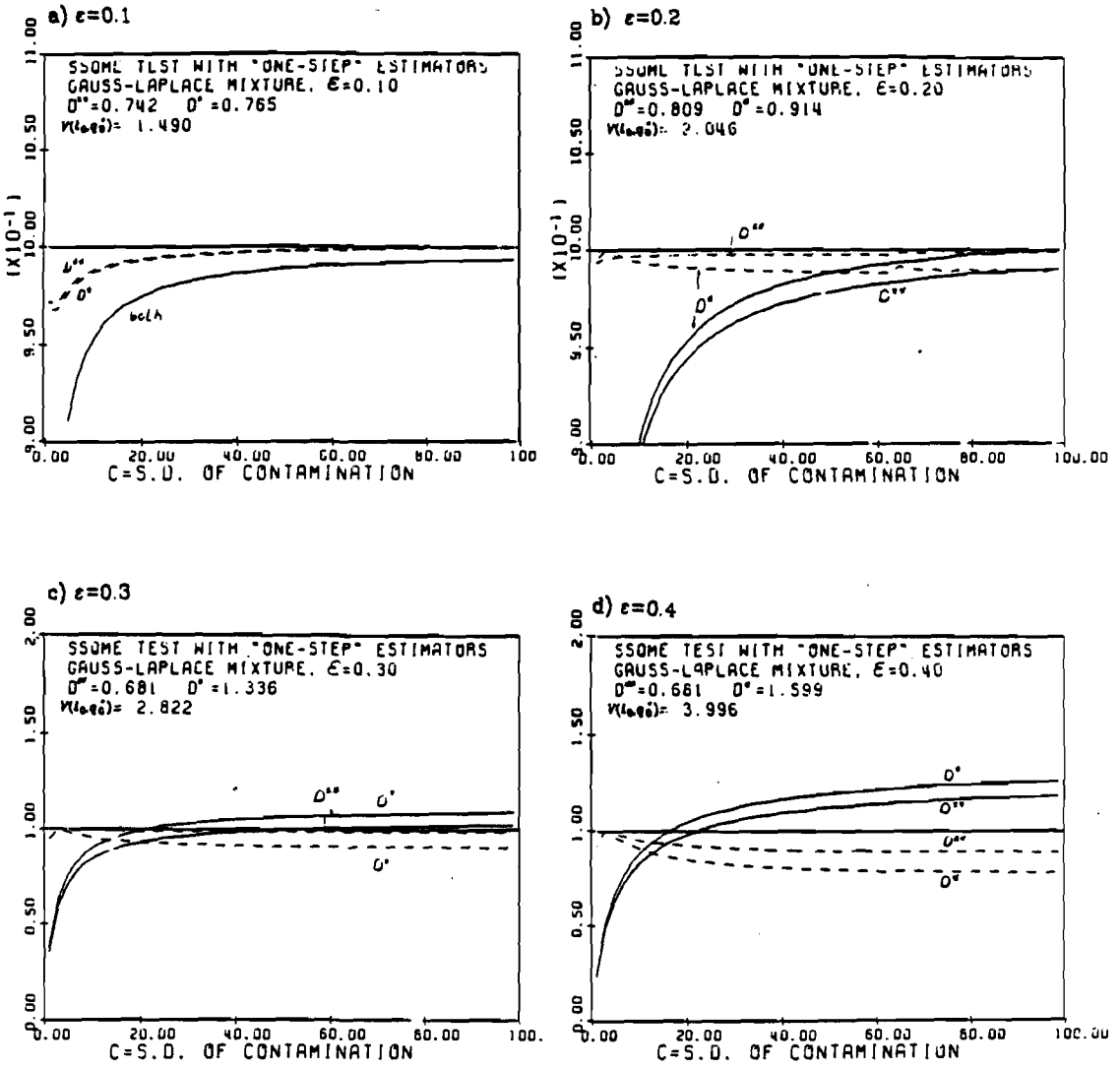


Fig. 5.4 Normalized variance $V'(F, T^{(1)})/V^{**}$ (dashed lines) and effective SNR loss $L(F, D)$ of SSQME test with a preliminary scale estimate. Gauss- Laplace mixture, c^2=contamination variance. D^* - scale normalization of the known-σ least favorable density, D^{**} - optimal scale normalization.

Finally, we note that the undesired result where the false alarm probability still depends on $f \in \mathbf{P}_0$ (unlike the SW-SSQME version) is a consequence of estimating the scale of the input samples and utilizing it for the threshold, instead of direct estimation of the variance of the location estimator. (Recall that in the classical normal case the variance of the sample mean estimator differs only by a factor of n from the variance of the input samples). This might be resolved with additional implementation complexity, by taking the adaptive threshold to be proportional to $\hat{V}^2(\mathbf{x})+\hat{V}^2(\mathbf{y})$, where

$$\hat{V}^2(\mathbf{x}) \;=\; \frac{1}{n}\left[\frac{1}{n}\sum_{i=1}^{n} IC^2\left(\frac{x_i - T_n^{(1)}(\mathbf{x})}{S_n^{(0)}(\mathbf{x})}, F_n(\mathbf{x}), T_n^{(1)}\right)\right] \tag{5.27}$$

and the IC is given by Eq. (5.13). Under the previously stated regularity and symmetry conditions this estimate of $V(F,T^{(1)})$ is consistent. Thus it is reasonable to assume that $V'\ (F,T^{(1)})$ will cancel out from the error probabilities. However, the error probabilities with this estimate are analytically intractable even when $n \rightarrow \infty$ (see also [3], section 6.8).

We feel that this additional complication is probably unnecessary, as is evident from Figs. 5.3-5.4. Moreover, Monte-Carlo finite sample simulation results indicate that when the test of Eq.(5.16) is implemented with the breakpoints corresponding to the density q_0^*, both error probabilities did not deviate in any noticeable way from the known σ case (i.e., when (5.12) is computed with $S(F)\doteq 1$), when $\epsilon \leq 0.2$ and $n > 50$.

5.3 Unknown Signal Frequency and Implementation Complexity

In radar-sonar detection, the frequency of the signal is often unknown, since it is proportional to the target radial velocity by the Doppler shift effect. In the Gaussian noise case, the problem is commonly solved by employing a bank of M contiguous coherent envelope tests , each of the form of Eq.(2.4), where the i^{th} is matched (by baseband conversion) to a Doppler shift of $i\ f_r/M$ ($f_r = 1/T$ =pulse repetition frequency). If the Doppler shift is spread uniformly over $[0,f_r]$, almost optimal performance over this range is obtained by implementing the contiguous bank with a weighted FFT processor of length n, as the envelope of the FFT filters outputs is almost flat.[1] This will serve in the following as a basis for comparison of implementation complexity of the various robust tests.

In a similar manner, robust tests for unknown frequency can be constructed as a bank of M contiguous SSQME tests. Specifically, in the k^{th} channel, $k=0,1,...M-1$, the input observables of Eqs.(2.5-2.6) are transformed according to

$$I_i(k)=Re\,[(x_i+jy_i)\exp(-j\,2\pi ik/M)]\ ,\ Q_i(k)=Im\,[(x_i+jy_i)\exp(-j\,2\pi ik/M)]$$

and then M parallel SSQME tests are performed on $\{\mathbf{I}(k),\mathbf{Q}(k)\}$, see Fig. 5.4a. In this way, for the M discrete frequencies kf_r/M in the uncertainty range $[0,f_r]$, the performance will be optimal as before, with some loss for straddling frequencies. The question now arises if this loss is comparable to that of the FFT processor when $M=n$. At first glance we might suspect that it would be larger due to the nonlinear processing, but it turns to be identical at least for large sample sizes.

Proposition5.3 Let the frequency shift of the signal relative to the down conversion mixers be f_d and let $n\rightarrow\infty$. Then, the detection probability of the various SSQME tests is given by the same expressions as for $f_d = 0$, where the effective SNR is attenuated by $G(f_d)$ which is identical to that of the FFT processor:

[1]Flatness of about 1 dB is achieved with Hamming weighting.

$$G(f_d) = \frac{\sin^2(n\pi f_d/f_r)}{\sin^2(\pi f_d/f_r)} \tag{5.28}$$

The proposition is proved in appendix D; the key property is an asymptotic normality result for M-estimators of a sequence of r.v.'s with unequal means. The implications should be clear: not only that it is not necessary to construct a bank of more than n contiguous SSQME tests, but it is also possible to apply the same weighting techniques that are employed with the FFT processor to get a flatter performance over the frequency range (at a price of reduced resolution and reduced SNR at $f_d=0$). (See appendix D for the details.)

The major increase in complexity stems from the nonlinear processing of the SSQME test. Thus, the number of operations required for a single frequency must be multiplied by n resulting in $O(n^2)$ for the full range, compared to $n\log n$ for the linear FFT processor (in the following, $\log n = \log_2 n$). It does not seem possible to perform some of the nonlinear operations *before* the n frequency conversions.

For purposes of simplicity, we assume that multiplications, additions and comparisons are equally costly. Hence, the FFT processor requires $O(5n\log n + 4n)$ real operations. The SSQME test for the nominally Gaussian noise is based on α-trimmed means. If ordering the data is done by the QUICKSORT algorithm [39], the expected number of operations for an i.i.d. sequence is $O(n\log n)$ [2]. Hence α- trimming requires $O(n\log n + c(\alpha), 1 \le c(\alpha) \le n$. With it available, the MAD scale estimator requires $O(n\log n + 2n)$ operations in addition, and the 1-step estimator $T^{(1)}$ for the non-Gaussian nominal density needs $O(11n)$ more, assuming that the evaluation of a nonlinear function is done by interpolation between two stored values. With these, the order of complexity for the various tests of the previous section is easily obtained by counting. For the SW version we do not count the operations in the adjacent M reference cells, as the SW is performed by sequentially repeating the algorithm on adjacent spatial cells and it

[2] QUICKSORT is a random algorithm, it can require up to $O(n^2)$ operations.

is only required to keep in memory the outcomes of operations that have already been performed. Table 5.4 summarizes the order of the number of operations. The last entry belongs to the narrow-band Wilcoxon nonparametric test [40] which is a candidate for comparison with the robust tests. The test statistic for a single frequency is

$$R_{NBW}(\mathbf{x},y) = [\sum_{i=1}^{n} R^{+}(x_i)U(x_i)]^2 + [\sum_{i=1}^{n} R^{+}(y_i)U(y_i)]^2 \tag{5.29}$$

where $R^{+}(x_i)$ is the rank of $|x_i|$ in $|x_1|, |x_2|, \ldots |x_n|$, and $U(x_i)$ is the unity step function. The test has the same order of complexity as the SSQME test, which substantially outperforms it as will be shown in the next section.

Type of Test	Order of Complexity
FFT processor	$5n\log n + 4n$
SW-SSQME for nominal Gaussian noise	$2n^2\log n + 2n(c(\alpha)+1)$
SSQME with preliminary scale estimate for nominal Gaussian noise	$4n^2(\log n + 1) + 2n(c(\alpha)+2)$
SW-SSQME for nominal non-Gaussian noise	$4n^2(\log n + 12) + 8n$
SSQME with preliminary scale estimate for nominal non-Gaussian noise	$4n^2(\log n + 12) + 11n$
Wilcoxon NB	$2n^2(\log n + 3) + 4n$

Table 5.4 Order of complexity for various tests of narrow-band input of length n.

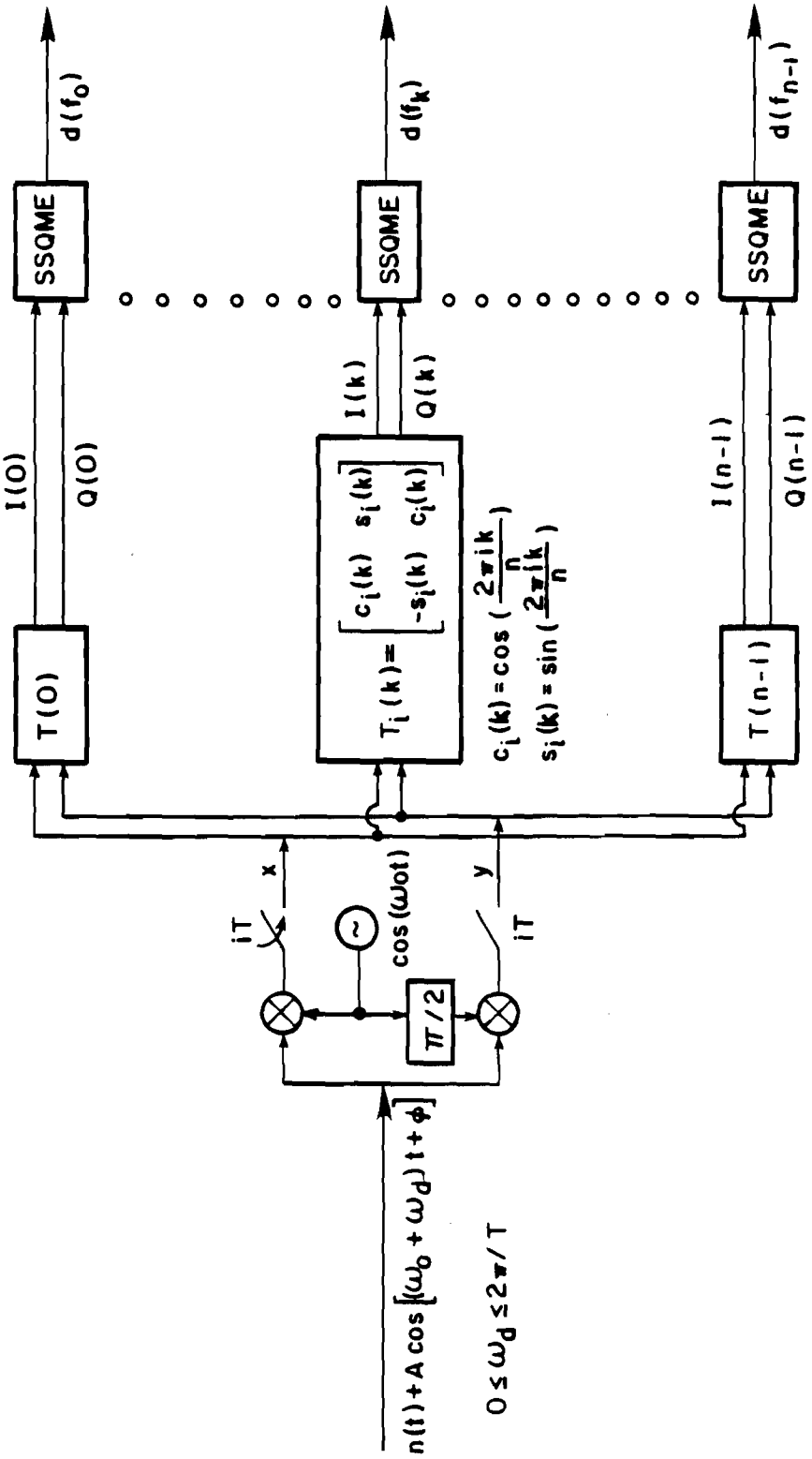


Figure 5.4*a* Bank of n SSQME tests for unknown Doppler frequency.

5.4 Finite Sample Simulation Results

The asymptotic approach to robust statistical procedures is an unavoidable necessity, as the $n \rightarrow \infty$ assumption allows one to invoke the central limit theorem in order to obtain *explicit* functionals which parametrize the performance of the procedure. Thus, optimization by some criterion becomes possible. Although asymptotically one could find true optimal procedures by estimating the underlying distribution, this usually requires a prohibitive number of samples [49]. Thus, the central reason for the asymptotic formulation is the hope that robust estimation and testing procedures will approach their asymptotic behavior quickly. The "Princeton Monte Carlo Study" of Andrews et.al. [35] showed that this is indeed the case for the variance of M and other robust estimators of location. For detection, we need also to demonstrate the tail behavior.

In most communication engineering applications, n will be quite small. This is true particularly for radar - sonar systems, where such considerations as the desired non -ambiguous detection range, angular resolution and minimization of the required time to search the detection spatial sector, call for decreasing n. Moreover, we have seen time and again through this work that detectability of *coherent* signals is governed (at least asymptotically) by the integrated SNR nA^2. In active systems, this product is proportional to the *average* transmitter power, which usually is the fundamental constraint rather than the peak power A^2. Therefore, at a given time period the same detectability can be maintained by increasing A and reducing n while keeping nA^2 constant, thus achieving the other important systems objectives. As a result of that, most radar systems are designed with $n = O(10\text{–}50)$ or even smaller for long range systems. This simple observation which is probably known to any practicing engineer has been totally overlooked in theoretical considerations, and the usefulness of procedures resulting from asymptotic analysis remains questionable without small - sample performance analysis.

Unfortunately, analytic tools for studying finite sample performance are not available. Results in the spirit of the Berry-Essen theorem [37], which bounds the deviation

from Gaussianity of the distribution of the sample mean estimator are non-existent for M-estimators. This is so because it is not possible to obtain an analytic expression for the finite sample moments. (Notice that its asymptotic variance was derived indirectly through the convergence of its distribution). We must resort to either numerical or simulation studies which do not analytically exhibit the rate of convergence.

For M-estimators with a monotone nonlinearity l, the distribution can be computed from an n-fold numerical convolution since $Prob\{\hat{A} < a\}$ $= Prob\{\sum l(x_i - a) < 0\}$. For our SSQME test more integrations are required since $R = \hat{A}_x^2 + \hat{A}_y^2$, and A and ϕ have to be averaged out. Field and Hampell [42] recently showed how to alleviate the enormous amount of computations by their so called "small sample asymptotics" technique, which is closely related to Daniels [43] saddlepoint approximation. Though this technique was shown to be very accurate down to sample sizes of 3, even in the extreme tails, it is still numerical in nature and involves a substantial programming effort. Therefore we have not proceeded in this direction but performed instead a Monte-Carlo study - programming is relatively straightforward and flexible, and the empirical distribution estimates are known to be unbiased and converge to the true distribution with a large number of repetitions.

Heuristically, the SSQME tests based on robust amplitude estimation should converge at a faster rate than the locally optimal robust detector (section 4.4) based on weak signal assumptions. The latter incorporates a hard "non tracking" nonlinearity; the distribution of the test statistic has finite support (when n is finite) with a point mass at the upper end of the interval. In contrast, the M-estimator can take any value on $\mathbf{R}^1$ and its distribution will be smoother and hence converge faster to Gaussian. This is most significant under H_1 when the distribution of the LORD test statistic will be mostly concentrated around the boundary of the intervals for large SNR.

When adopting the asymptotic theory testing structure for finite sample size, the amount of trimming or limiting must be increased if very small P_{fa} is desired, since we

are dealing with the tail of the distribution where deviations from theory are larger. This is explained more precisely as follows. Define A_k = { *number of outliers in either of the quadrature channels that is exactly k, and less than or equal k in the other channel* }. Since these events are mutually exclusive and their union is the certain event we can write

$$P_{fa} = Pr\{R \geq t\} = \sum_{k=0}^{2\lfloor n\alpha \rfloor} Pr\{R \geq t \mid A_k\} P(A_k) + \sum_{k=2\lfloor n\alpha \rfloor+1}^{n} Pr\{R \geq t \mid A_k\} P(A_k) \tag{5.30}$$

Thus, a lower bound for P_{fa} is the second term. In it, the number of outliers exceeds the trimming capability of either estimator; hence, if t is fixed according to asymptotic theory $Pr\{R \geq t \mid A_k\}$ will be large and close to 1, and approximately $P_{fa} \cong \sum_{k \geq 2\lfloor n\alpha \rfloor+1} P(A_k)$. Assuming that the contaminated samples come from a switching model, the probability of having k outliers out of n in either channel is binomial: $P_k = \binom{n}{k}(1-\epsilon)^{n-k}\epsilon^k$, and due to independence of the quadrature samples

$$P(A_k) = P_k^2 + 2P_k \sum_{l=0}^{k-1} P_l \tag{5.31}$$

The following table exhibits this approximation for $n=16$ and $\epsilon=0.1$. With the asymptotically optimal $\alpha=0.164$, P_{fa} can not be smaller than roughly 10^{-3}, while $\alpha=0.27$ is sufficient for 10^{-6}; simulation results have validated this approximation.

α	.125	.1875	.25	.3125	.375
$k_m = 2\lfloor n\alpha \rfloor + 1$	5	7	9	11	13
$\sum_{k=k_m}^{n} P(A_k)$	$3.37 \cdot 10^{-2}$	$1.01 \cdot 10^{-3}$	$1.18 \cdot 10^{-5}$	$5.41 \cdot 10^{-8}$	$8.36 \cdot 10^{-11}$

Several versions of scale-invariant detectors of section 5.2 were studied, for the nominally Gaussian case. They will be denoted D1 - D5 as follows:

(D1) SW-SSQME test of Eq.(5.5) based on α– trimmed estimators with the number of reference cells $M=4$.

(D2) like D1 but $M \rightarrow \infty$ (i.e.,the estimation of the variance of the test statistic converges to the true value). For that, the reference statistics $W(\mathbf{u},\mathbf{v})$ of the D1 simulation was accumulated over all runs to produce the adaptive threshold. Thus, a huge saving was possible.

(D3) SW-SSQME test with a "one-step" location estimators of Eq.(5.15) based on the MAD scale estimate; its normalization factor D=0.632 was adjusted to get (empirically) consistency of the scale estimator $ES_{16}^{(0)}(\mathbf{x})=1$ at the Gaussian (this value is somewhat lower than predicted from the asymptotic theory). The nonlinearity was the soft limiter $l(x,-k,k)$.

(D4) SSQME test of Eqs. (5.16)-(5.17) - no reference cells - based on the same location and scale estimators of D3 .

(D5) SW version of the LORD from section 4.4 (obtained from weak signal assumptions by replacing the M - estimators with the nonlinearity-integrator of Eq. (4.29)), $M \rightarrow \infty$.

For D1-D4, both α and k were adjusted to obtain better false-alarm control as discussed before; for D5 we chose k as the optimal for the given ϵ from the asymptotic theory.

For each test, two simulation programs were built. The first is for H_0 and its output are graphs of the false alarm probability vs. the normalized threshold, as well as the first two moments of the location and scale estimators. The second program is for H_1. It simulates the detection probability as a function of the SNR for several levels of the false alarm probability and for a Swerling I target model, i.e., coherent narrowband signal with Rayleigh distributed amplitude and uniform phase (see section 4.3, case a). In each of the simulations, the test was subjected to four different noise p.d.f.'s from an ϵ-mixture family where the nominal is normal Gaussian:

(f_1) $\epsilon = 0$

(f_2) $\epsilon = 0.1$ and two point masses at ± 15

(f_3) $\epsilon = 0.1$ and a Gaussian contamination with variance $c^2=100$

(f_4) $\epsilon=0.01$ and a Gaussian contamination with $c^2=900$

In the sequel, we shall use these numbers to identify the situation. The input variances are thus 1, 23.4, 10.9 and 9.99 respectively - all $\epsilon \neq 0$ cases represent large, heavy tailed contaminations and would have resulted in a substantial degradation of performance if the unrobustified envelope detector was employed (see Section 2.2). The ϵ– mixture r.v.'s were generated from a switching device: at each step, a normal r.v y_i was generated together with a binomial r.v s_i that is equal to 1 with probability ϵ; $x_i = y_i$ if $s_i \neq 1$ and $x_i = c \cdot y_i$ (or $x_i = \pm 15$ with probability 0.5 for f_2) otherwise. Almost all simulations were for $n=16$.

Table 5.5 displays the Monte- Carlo variances of the α– trimmed estimator for various α's. For comparison, from the asymptotic theory at the least-favorable p.d.f. $V^*(\epsilon=0.1) = 1.49$, and for $\alpha=0.225$ we computed from Eq.(5.3) $V(1)=1.17$, $V(2)=1.51$, $V(3)=1.46$, $V(4)=1.2$. This good correspondence between asymptotic theory and finite sample simulation (for n=20) was observed before by Andrews et. al. [35]. The asymptotically optimal trimming ratio is $\alpha(\epsilon=0.1)= 0.164$; as was discussed before it is necessary to apply more trimming with small sample size when very small false alarm probabilities are desired. The simulations indicated that $\alpha=0.225$ to 0.3 is sufficient; the table shows that except for f_1 for this α range the variances are roughly as those for $\alpha = 0.164$ (or even smaller). The results suggest that the detectability loss compared to theory is small. This has been verified from the H_1 simulation results, which will be presented below. Similar conclusions can be drawn from Table 5.6, which is for the "one-step" estimator; here the asymptotically optimal limiting is $k(\epsilon=0.1) = 1.14$, and very small losses (if any) are incurred by taking $k=0.75$.

α	$V(f_1)$	$V(f_2)$	$V(f_3)$	$V(f_4)$
0.164	1.09	1.43	1.60	1.13
0.225	1.15	1.34	1.53	1.18
0.25	1.17	1.35	1.51	1.20
0.3	1.21	1.41	1.55	1.25
0.4	1.36	1.56	1.69	1.39
0.5	1.47	1.68	1.81	1.50

Table 5.5 Monte-Carlo variances of the α– trimmed estimator. n=16, 20,000 runs.

k	$V(f_1)$	$V(f_2)$	$V(f_3)$	$V(f_4)$
1.14	1.08	1.29	1.53	1.12
0.9	1.13	1.31	1.51	1.16
0.75	1.16	1.34	1.50	1.24
0.5	1.29	1.46	1.62	1.32
0.3	1.41	1.60	1.74	1.42

Table 5.6 Monte-Carlo variances of the "one - step" estimator with MAD scale estimator. n=16, 20,000 runs.

Probability of false alarm curves are shown in Figures 5.5 - 5.16. In all these figures, the smooth curves marked "T" are computed from the asymptotic theory - Eqs.(5.6) and (5.23), and the numbered curves are the simulation results for the densities f_1– f_4. Fig. 5.5 corresponds to D1 and 5.6 to D2, with α=0.165. We see clearly that here α is too large, since for the more contaminated cases P_{fa} is much higher than desired. This problem is solved by taking α=0.225 as can be seen from Figs. 5.7-5.8. In the first one, the Monte-Carlo curves are all **remarkably close** to the theoretic curve, within the range where the simulation results are reliable [1], $P_{fa} \gtrsim 4/n_r = 2\cdot 10^{-4}$. The test is clearly CFAR, for very different noise p.d.f.'s. In Figure 5.8 P_{fa} is even somewhat smaller than the theoretic, (this means that the distribution of the test statistic R has shorter tails compared to the one sided exponential). Thus, if the threshold is fixed according to the asymptotic theory, P_{fa} is bounded from above as desired. The same is valid for α=0.3 as can be seen from Figs. 5.9-10. The curves for D3 are shown in Figs.

[1] The empirical distribution for n_r repetition of the Monte-Carlo experiment $\hat{F}_{n_r}(x) = \frac{1}{n_r}\text{number}\{x_i \geq x\}$, is a binomial r.v. with mean$=1-F(x)$ and variance $=\frac{1}{n_r}F(x)(1-F(x))$. Therefore, for $P_{fa} << 1$,for $\sqrt{Var(\hat{P}_{fa})}/P_{fa} \leq 0.5$ one obtains $P_{fa} \geq 4/n_r$.

5.11-12; in the first one, the limiter breakpoint is $k=1.14$ according to the asymptotic theory - it is again seen that the limiting is not "hard" enough, as P_{fa} for f_3 is higher than the theoretic bound when $P_{fa}<0.01$. This problem is corrected by taking $k=0.75$ as can be seen from Fig. 5.12.

Results for D4 are displayed in Figs. 5.13-5.14; notice the two sets of curves in Fig. 5.13: the first is of $\log Prob\{R(\mathbf{x},\mathbf{y})\geq(t/n)V^* W(\mathbf{x},\mathbf{y})\}$ where $V^*=1.49$, which should approach a straight line when $n\rightarrow\infty$ according to proposition 5.2. In the second case the scale is contracted by taking $V^*=3$. Here asymptotic theory is no longer valid: the curves are not linear and for achieving small P_{fa}'s higher threshold settings are required. These higher settings, in turn, would cause higher detectability losses. This deviation from asymptotic theory is not surprising, and is equivalent to the behavior of the "Quadrature t-test," which is the maximum likelihood test for a coherent narrowband signal in Gaussian noise of unknown variance

$$(\frac{1}{n}\sum x_i)^2+(\frac{1}{n}\sum y_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{t}{n}[\sum(x_i-\bar{x})^2+\sum(y_i-\bar{y})^2] \qquad (5.32)$$

For small n, the right hand side cannot be approximated by a Gaussian distribution (it is chi-squared *for all* n), and also the dependency between the right and left sides of Eq. (5.29) cannot be neglected as we did in proving Proposition 5.2. In Fig. 5.14 n was increased to 48, in order to check convergence. It is evident that the curves almost converge to asymptotic theory, although some curvature is still noticed.

Finally, Fig. 5.15 is for the LORD test D5. Here, as in Fig. 5.8, a tendency for shorter tails is observed. This could be heuristically described by the hard limiting that restricts the distribution of the test statistic to have a finite support.

Dramatic improvement in false alarms control over the unrobustified SW detector Eq. (2.8) is evident when comparing the previous figures with Fig. 5.16. When it operates in a Gauss-Gauss ($\epsilon=0.1$, $c=10$) noise, P_{fa} is intolerably increased from the design 10^{-4} to about 0.1 when $M=4$, and from 10^{-6} to $2\cdot10^{-4}$ when $M\rightarrow\infty$. Compare also Fig.

2.5 for ϵ=0.01. Thus we can confidently conclude that the various SW-SSQME tests essentially maintain CFAR according to the asymptotic theory of section 5.2 even for n as small as 16 with P_{fa} down to (at least) 10^{-4}. For the test without reference samples D4, n is not much larger than 48. Based on these curves, the detection thresholds for the H_1 simulation were simply taken from the smooth theoretical curves, and extrapolated from them for $P_{fa}=10^{-6}$. Future work could utilize reduced variance sampling techniques such as "importance sampling "[44] to validate this extrapolation.

Probability of detection curves for D1, D2 and D5 are shown in Figs. 5.17-27. Here 5000 repetitions are sufficient to get $Var^{1/2}(\hat{P}_d)/P_d \leq 0.1$ for $.02 \leq P_d \leq 0.98$. The ordinate is the effective integrated $SNR = nE(A^2)/\sigma^2$ where σ^2 is the variance of the *nominal* Gaussian p.d.f. The known signal frequency case is depicted in Figs. 5.17-20. The Monte-Carlo curves are clearly bounded from the left by the computed P_d for the unrobustified SW detector when ϵ=0, and from the right by the maximin lower bound computed from asymptotic theory, Eq. (5.6), with V^* of the least-favorable p.d.f. f^*. Here again the fit between asymptotic theory and small sample performance is striking: the curves for f_1–f_4 are ranked *exactly* according to the estimation variances of Table 5.5, and the simulated P_d values generally differ by no more than a few percent when the simulation variances are substituted in Eq. (5.6) or (5.23). Equivalently, the horizontal differences between curves are roughly the dB values of the variances from Table 5.5. We also note that the difference between α=0.225 and α=0.3 is very small, again according to the variance difference.

In contrast with the above results, the unrobustified SW detector of Eq.(2.8) was found to suffer substantial detectability losses. They were almost identical to the increased *input* variances of $f_2 - f_4$, with respect to f_1.

Some non-parametric schemes were developed for detection of narrowband signals in unknown noise; the narrowband Wilcoxon detector, Eq.(5.29), was proposed by Carlyle and its Monte-Carlo performance in purely Gaussian noise was studied by Hansen

[40]. It was found that the SNR losses compared to the UMP detector for n=16 and P_{fa}=10^{-6} are 5 and 14 dB, for P_d=0.5 and 0.9, respectively (smaller losses are incurred for larger n) - compared to merely 0.6 db for D2 in Fig. 5.18 ! Furthermore, the performance in non-Gaussian noise has not been studied.

The performance of the detector without reference samples, $D4$ with $n = 16$, is shown in Fig. 5.20. Although convergence to the asymptotic theory is not yet reached (in accordance with the higher required threshold values - recall Fig. 5.13), the detector is clearly robust. This is evident from the small differences between the curves for $f_1 - f_4$. By comparison with Fig. 5.18, the SNR losses are roughly 1.8 and 5.8 dB for $P_{fa} = 10^{-2}$ and 10^{-6}, respectively, independent of the P_d. However, when comparing it with the SW-SSQME test D1 of only 4 reference channels, D4 outperforms it by roughly 1 and 3.4 dB, for $P_{fa} = 10^{-2}$ and 10^{-6}, respectively. By virtue of the shift invariance of the one-step location estimator, if convergence to Gaussianity is reached under H_0, the performance under H_1 will agree with the asymptotic prediction. Thus, when $n \geq 50$ D4 would hardly suffer any detectability losses compared to the UMP detector for the least favorable noise p.d.f.

In contrast to the excellent performance of the structures proposed in this work, Fig. 5.21 shows that while the LORD test D5 performs roughly the same for P_{fa}=10^{-2}, it is *inferior* for P_{fa}=10^{-6}; its SNR should be higher by 4.6 dB to achieve P_d=0.9 at f_3. This can be simply explained: when the SNR is sufficiently high for high P_d, the "non-tracking" hard limiter of Eq. (4.29) prevents the distribution of R from tilting towards higher values. Also, asymptotic theory does not properly describe its behavior at the most interesting zone $P_d \rightarrow 1$. (Recall that SNR $\rightarrow 0$ was *necessary* for the derivation of its asymptotic distribution, but not for the tests which are based on true robust estimation of the arbitrary amplitude.)

The performance with unmatched signal frequency when a bank of n contiguous SW-SSQME tests is utilized for unknown frequency (section 5.3), are depicted in Figs.

5.22-24 where the normalized deviation $\Delta f \triangleq n \Delta f_d / f_r = 0.25$, and in Figs. 5.25-27 for $\Delta f = 0.5$ (halfway between the central frequency of adjacent tests). The mismatch loss is almost indistinguishable from that of Proposition 5.3 - i.e., the same loss incurred by the FFT processor for a frequency unmatched signal in pure Gaussian noise. While this loss is reasonable for D1 and D2, 0.91 dB and 3.9 dB for $\Delta f = 0.25$ and 0.5 respectively, the weak signal locally - optimal test D5 **totally breaks down** under this situation, as demonstrated by Figs. 5.24 and 5.27. We note that this loss for D_1 and D_2 can be reduced by weighting the input samples. (See appendix D.)

Another interesting conclusion can be drawn. In many practical cases, the noise p.d.f. is variance constrained, as it is proportional to the total (nominal plus contaminating components) noise power. This is true in particular for the clutter (reverberation) environment, where the noise power is proportional to the radar (sonar) transmitted power. A question now arises as to which variance constrained noise p.d.f. constitutes the worst detection environment. Consider for example Fig. 5.18 and f_3 for which the detector D2 is essentially optimal. However, the variance of f_3 is 10.9. Consequently, in a Gaussian noise environment of equal power $\sigma^2 = 10.9$, the UMP detector would perform worse by 10.9 -1.5 = 9.4 dB than the detector D2 ! This turns to be true in general. In [45], we show that the Gaussian is the least favorable variance constrained p.d.f. for detection of lowpass as well as narrowband coherent signals with random parameters.

Fig. 5.5 False alarm probability. Test - D1 , α=.164.

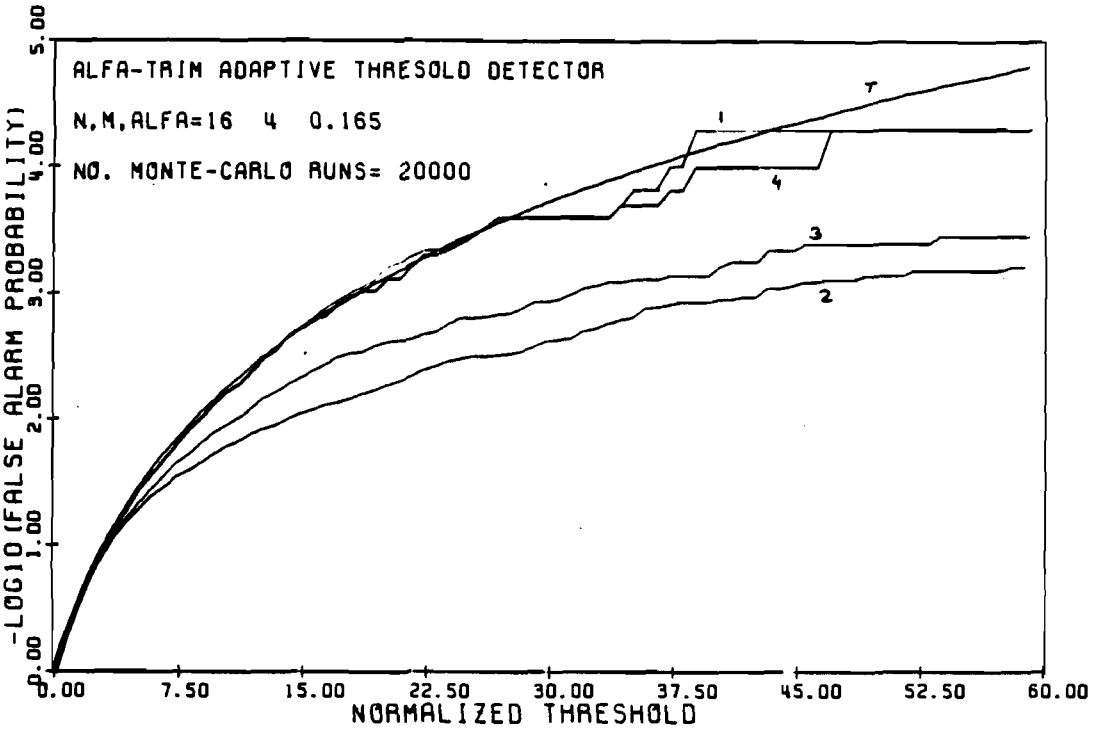


Fig. 5.6 False alarm probability. Test - D2 , α=.164.

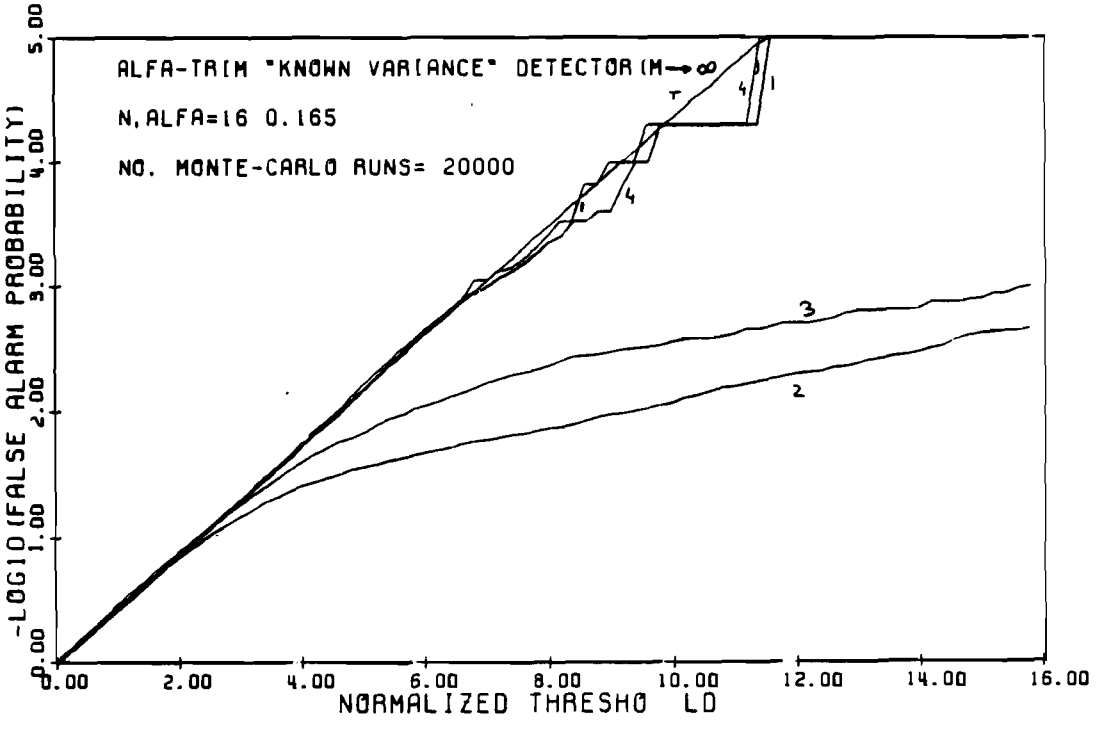


Fig. 5.7 False alarm probability. Test - D1 , α=.225.

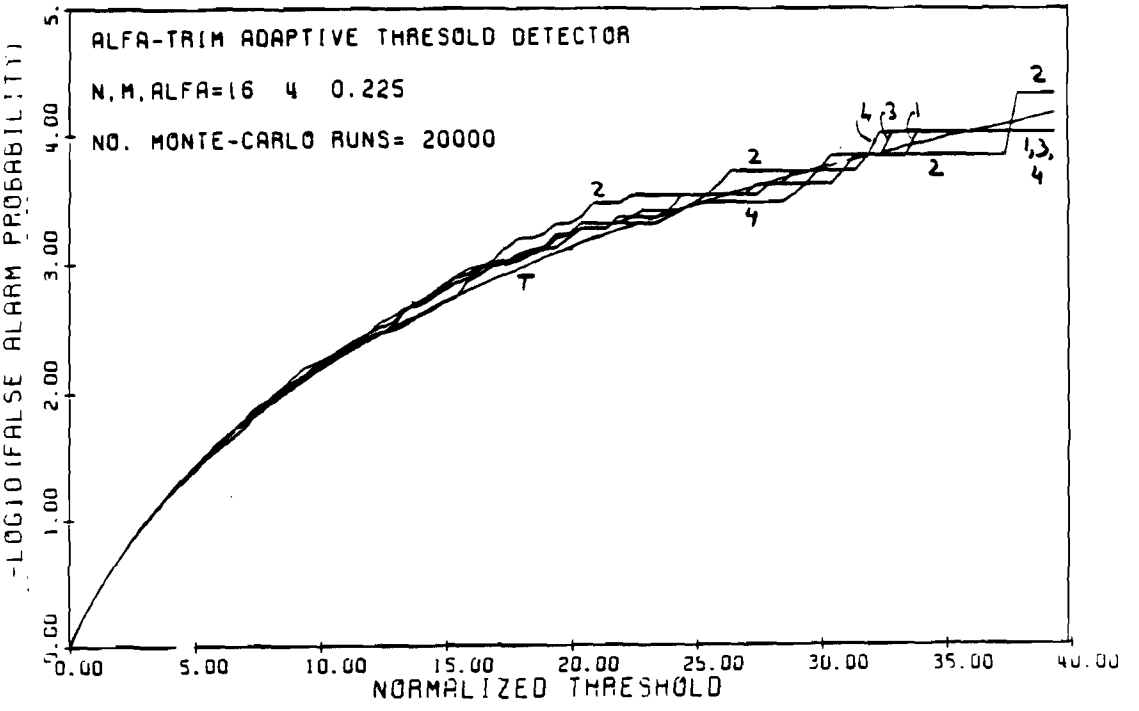


Fig. 5.8 False alarm probability. Test - D2 , α=.225.

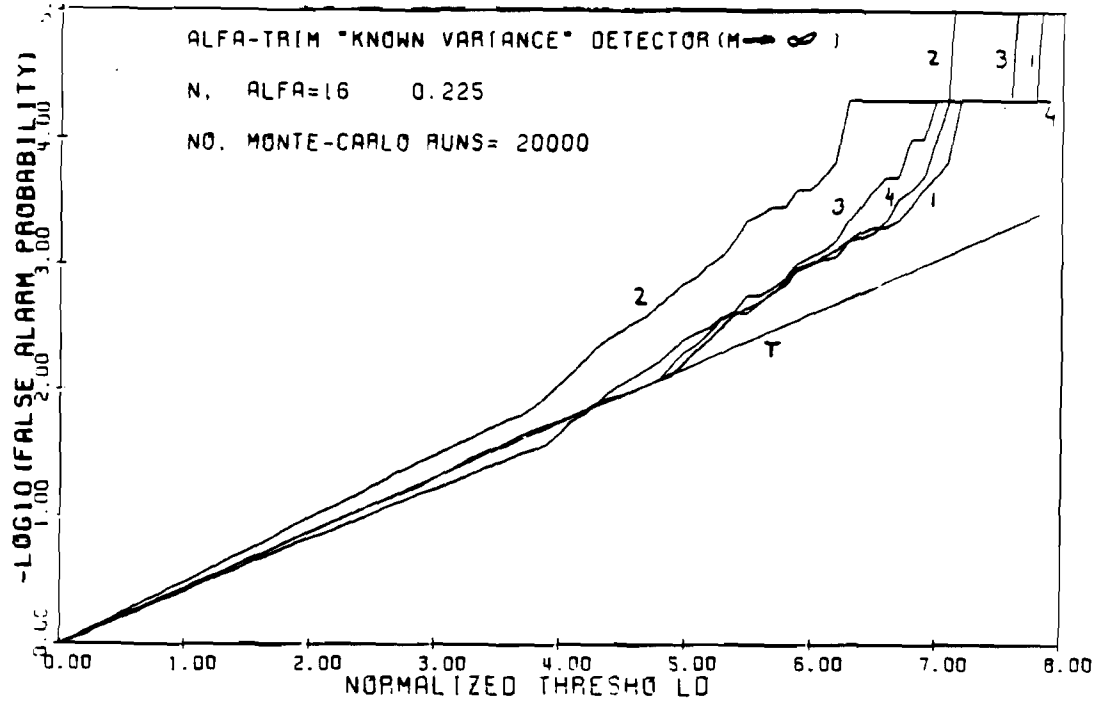


Fig. 5.9 False alarm probability. Test - D1 , α=.3.

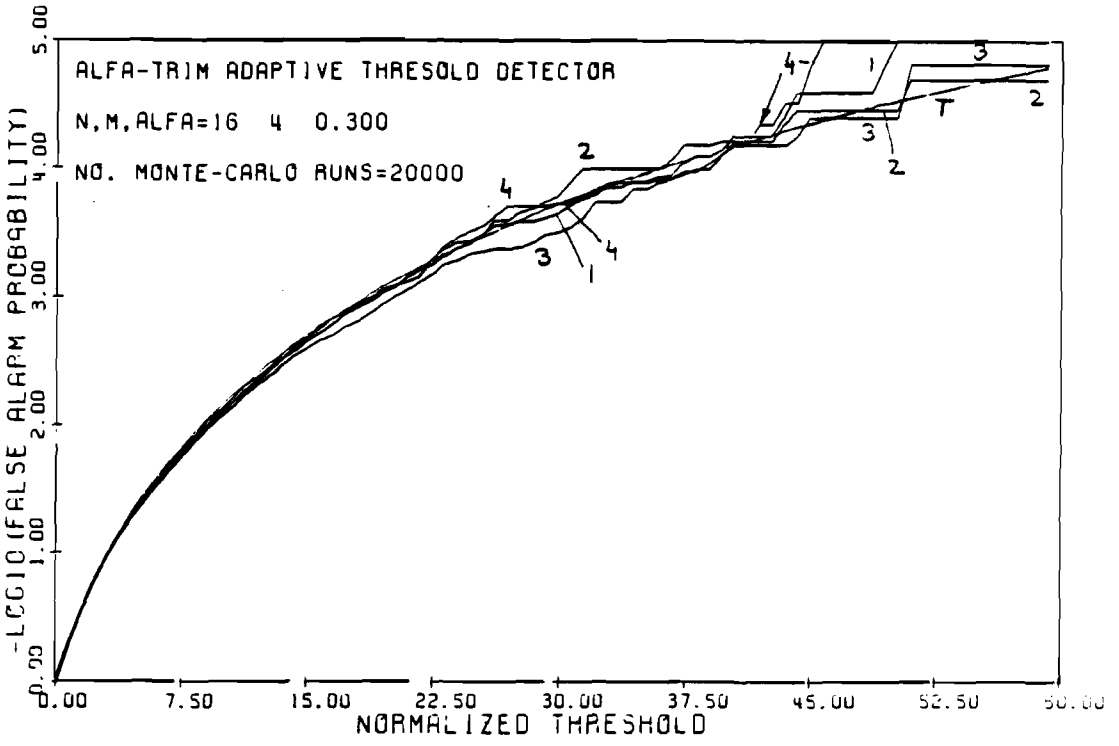


Fig. 5.10 False alarm probability. Test - D2 , α=.3.

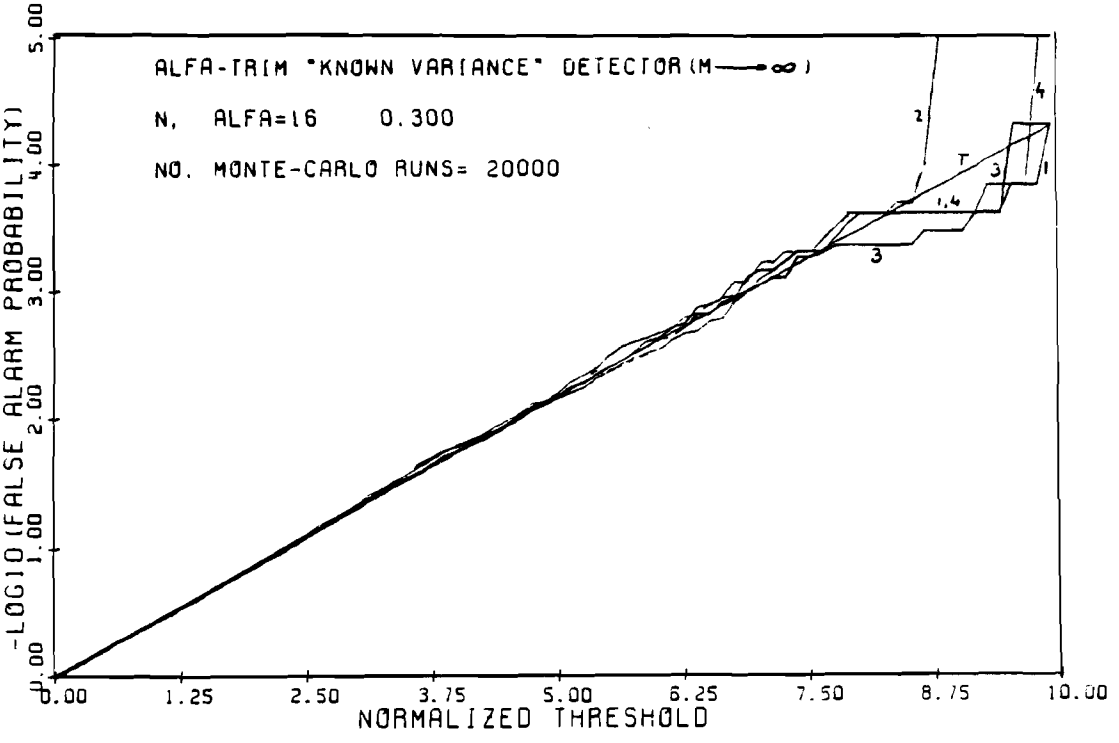


Fig. 5.11 False alarm probability. Test - D3 , K=1.14.

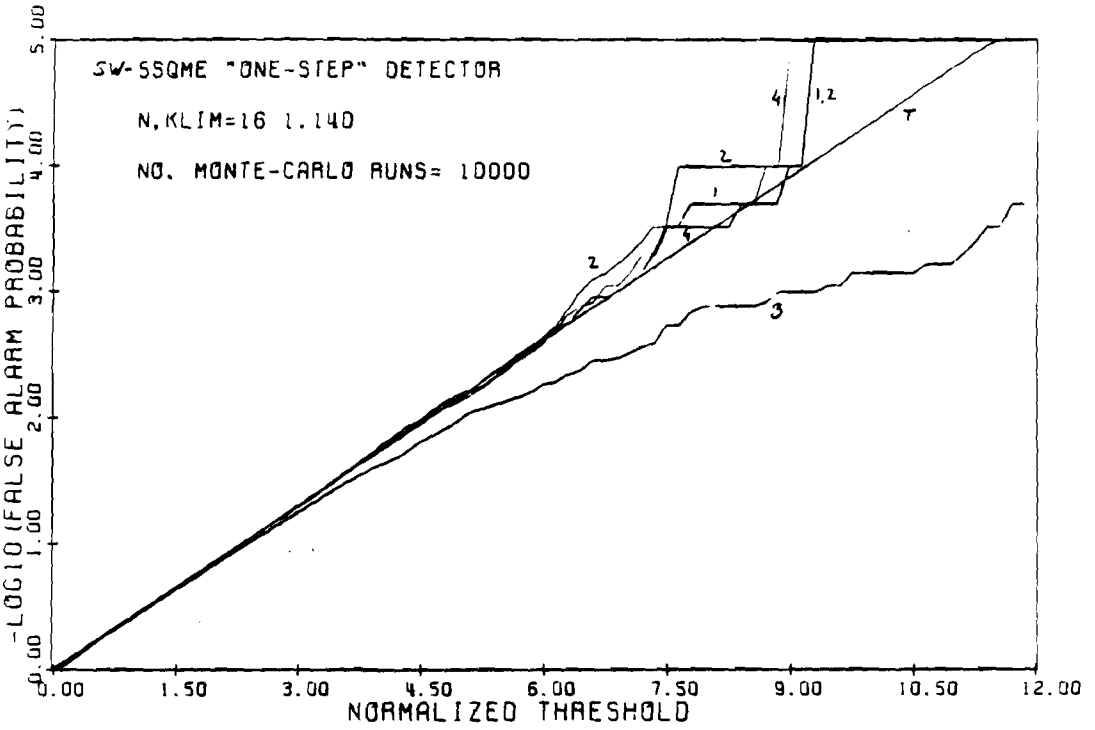


Fig. 5.12 False alarm probability. Test - D3 , K=.75.

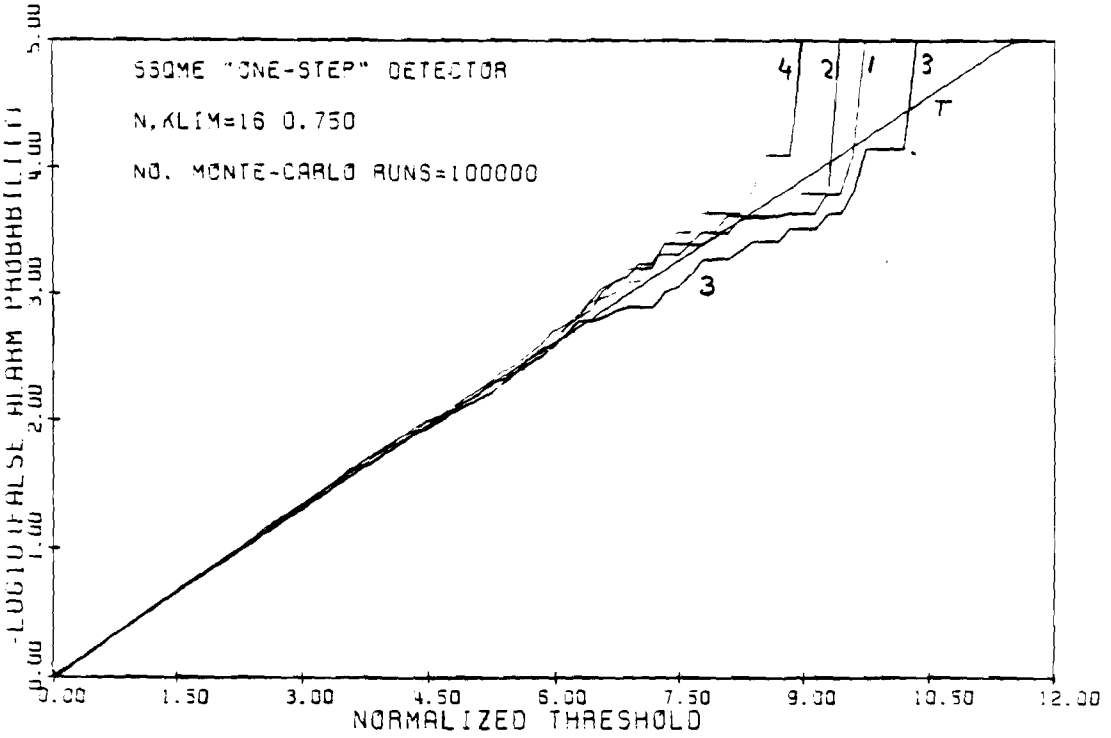


Fig. 5.13 False alarm probability. Test - D4 , K=.85, n=16.

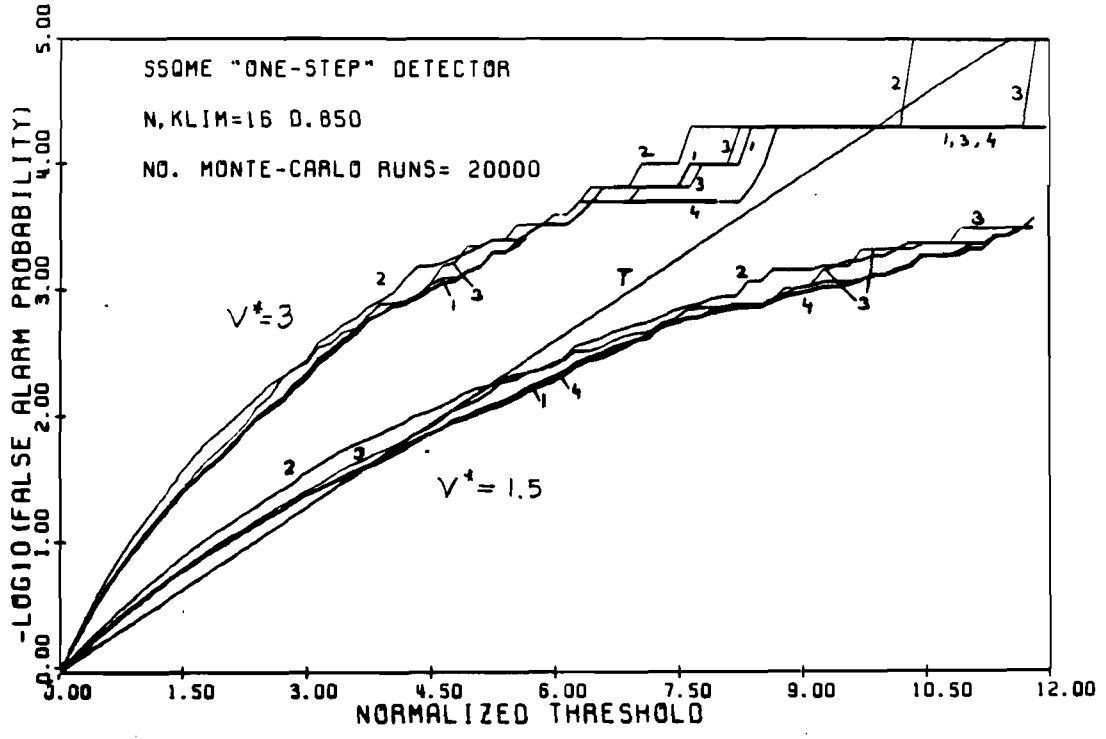


Fig. 5.14 False alarm probability. Test - D4 , K=.9, n=48.

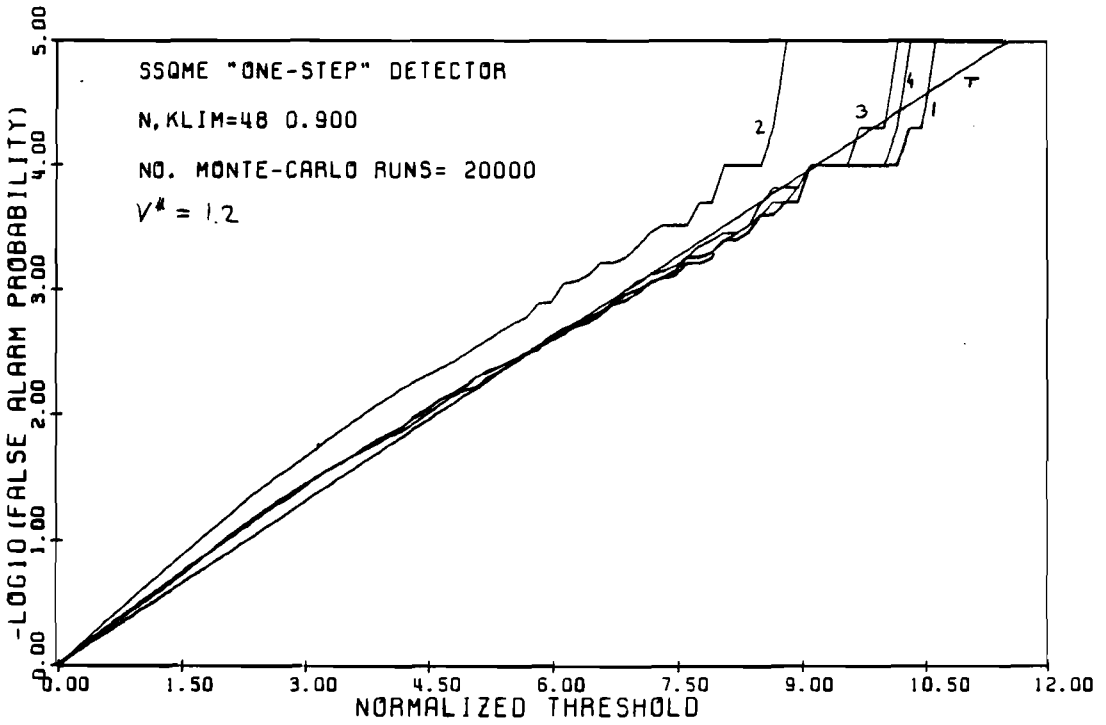


Fig. 5.15 False alarm probability. Test - D5 , K=1.14.

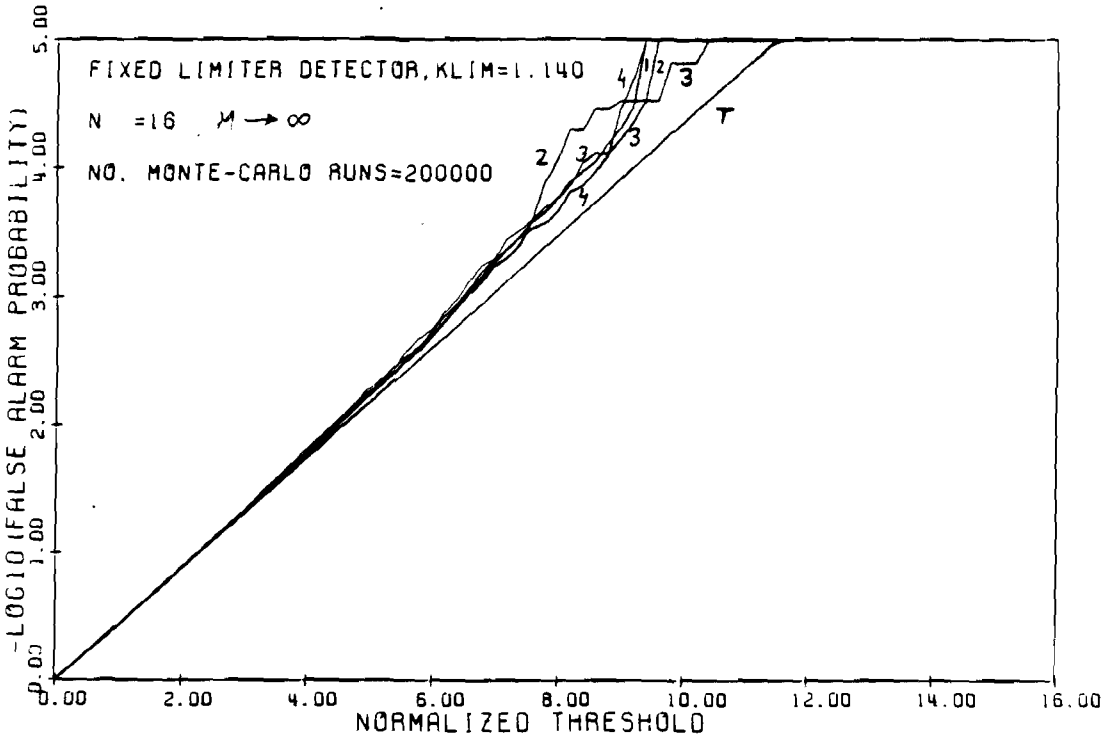


Fig. 5.16 False alarm probability. Unrobustified envelope detector, Gauss-Gauss mixture (ϵ=0.1, c =10).

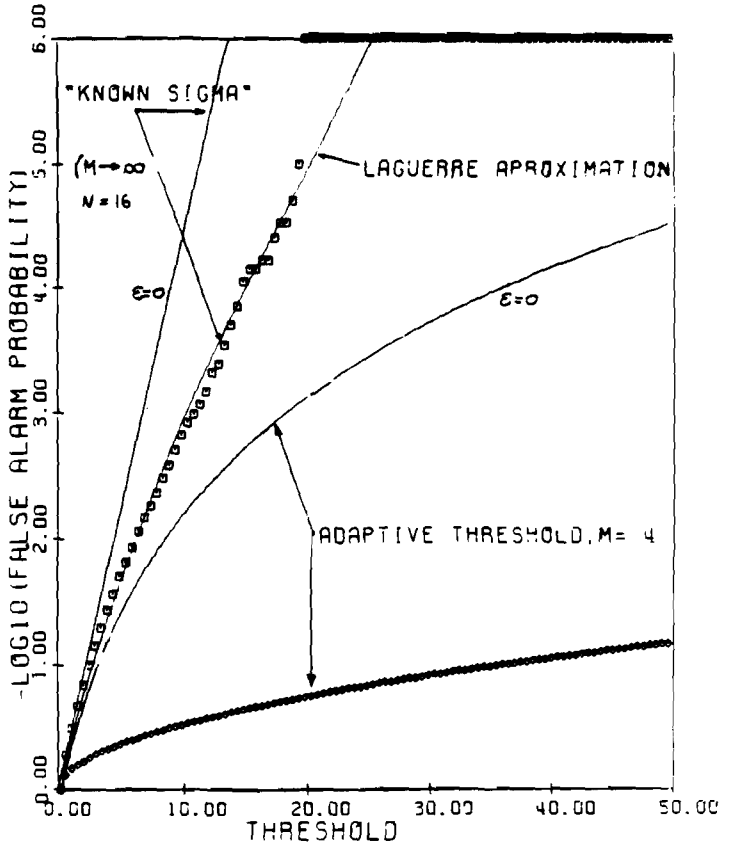


Fig. 5.17 Detection probability. Test - D1, α=.225,Δf =0

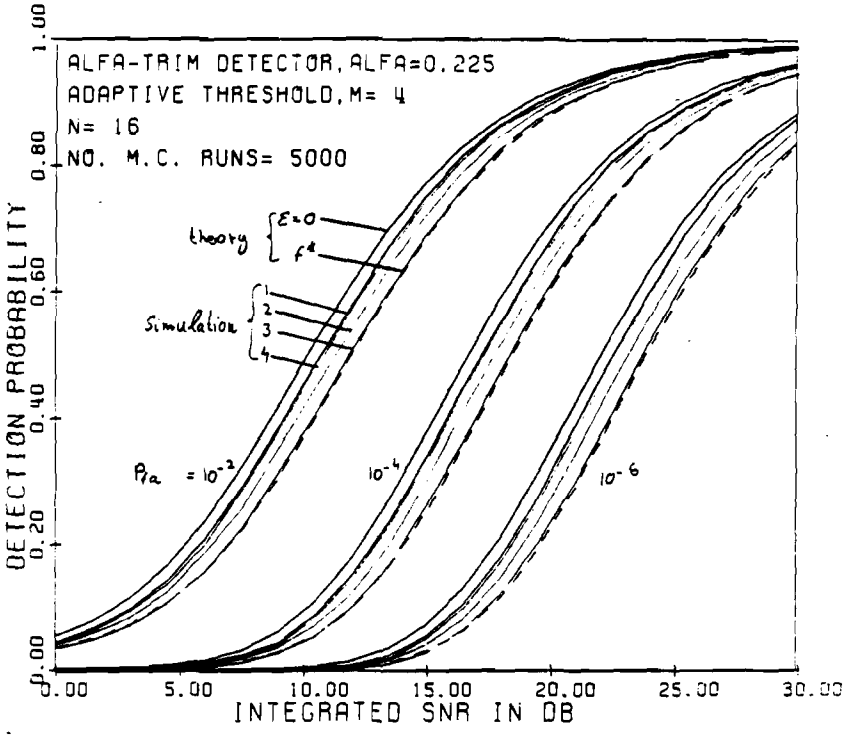


Fig. 5.18 Detection probability. Test - D2, α=.225,Δf =0

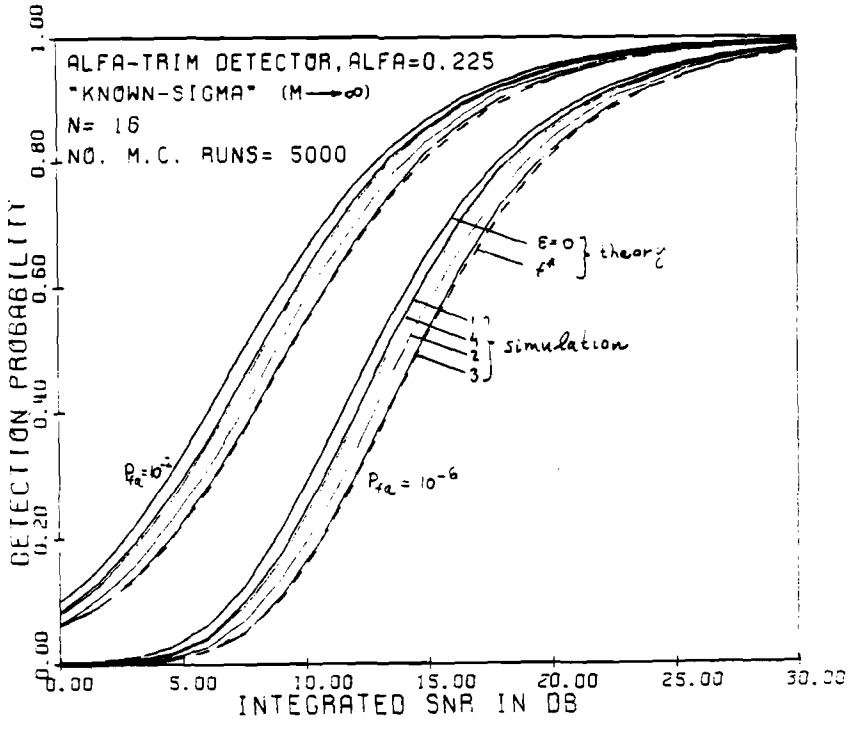


Fig. 5.19 Detection probability. Test - D1, α=.3,Δf =0

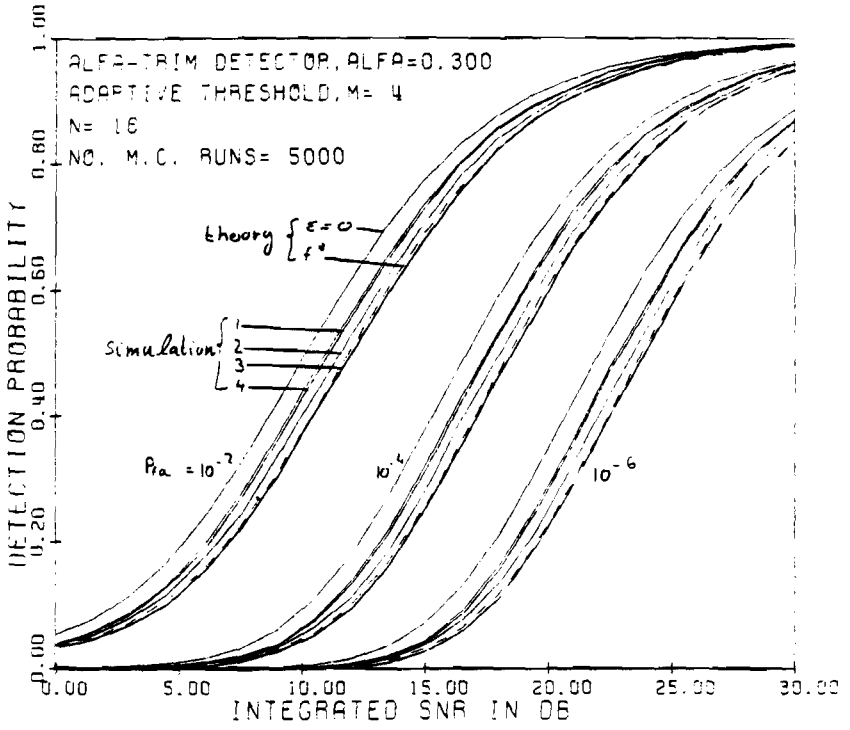


Fig. 5.20 Detection probability. Test - D4, n = 16, Δf =0

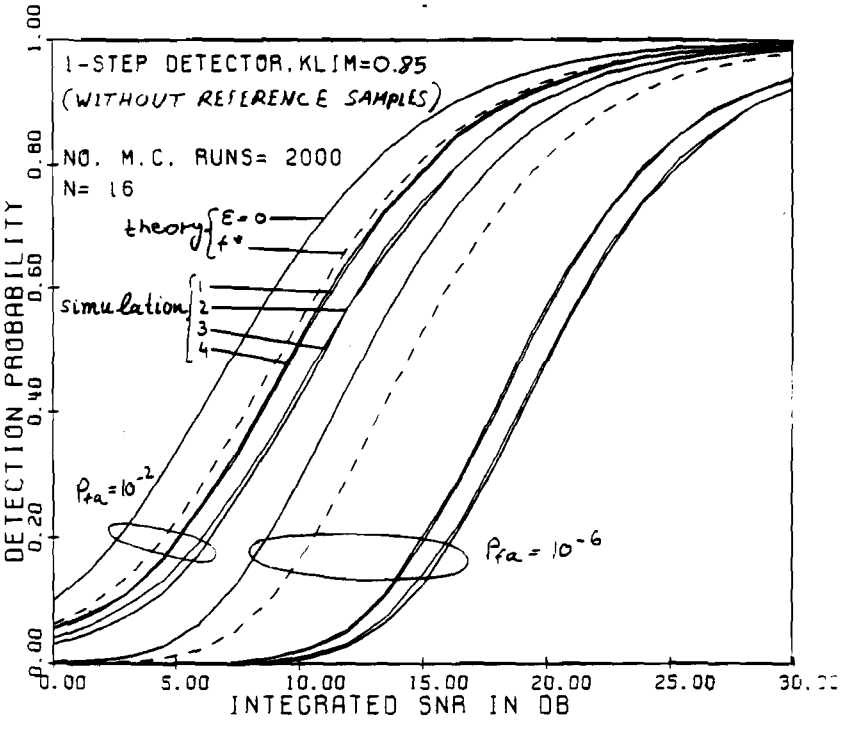


Fig. 5.21 Detection probability. Test - D5, K = 1.14, Δf =0

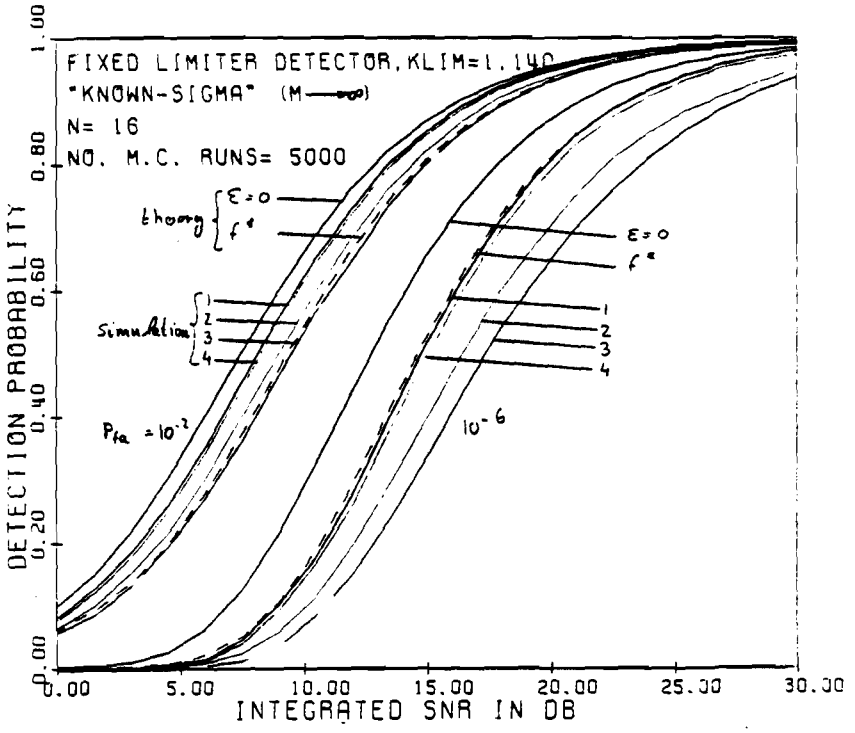


Fig. 5.22 Detection probability. Test - D1, α=.225,Δf =.25

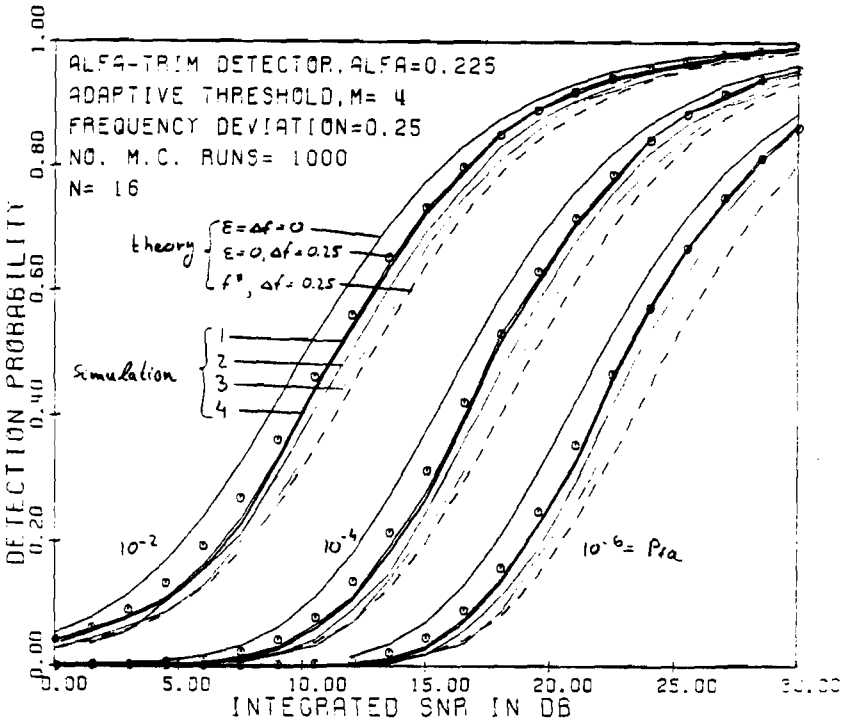


Fig. 5.23 Detection probability. Test - D2, α=.225,Δf =.25

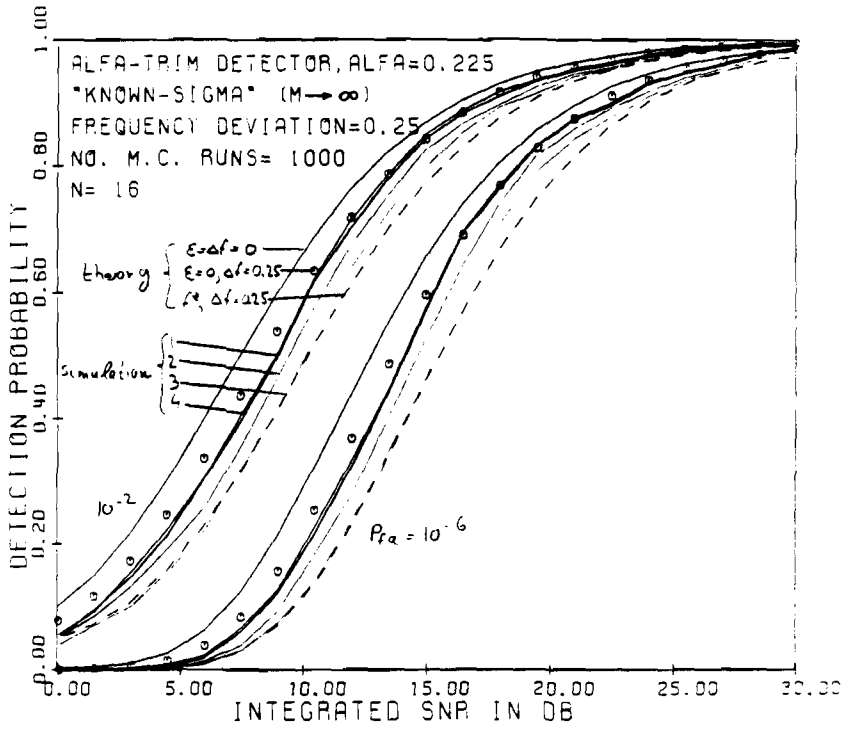


Fig. 5.24 Detection probability. Test - D5, K =1.14,Δf =.25

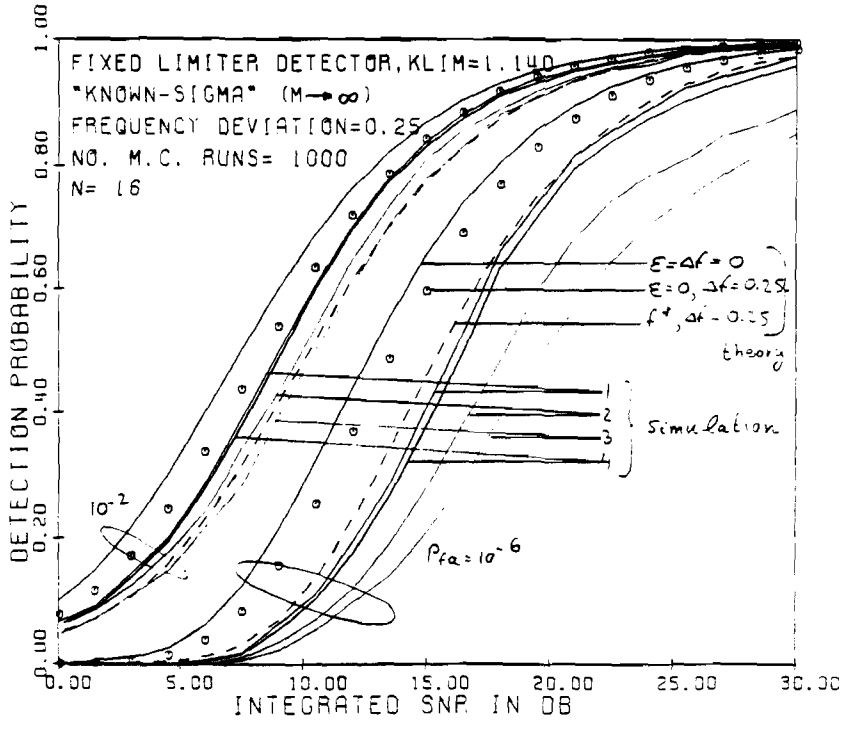


Fig. 5.25 Detection probability. Test - D1, α=.225,Δf =.5

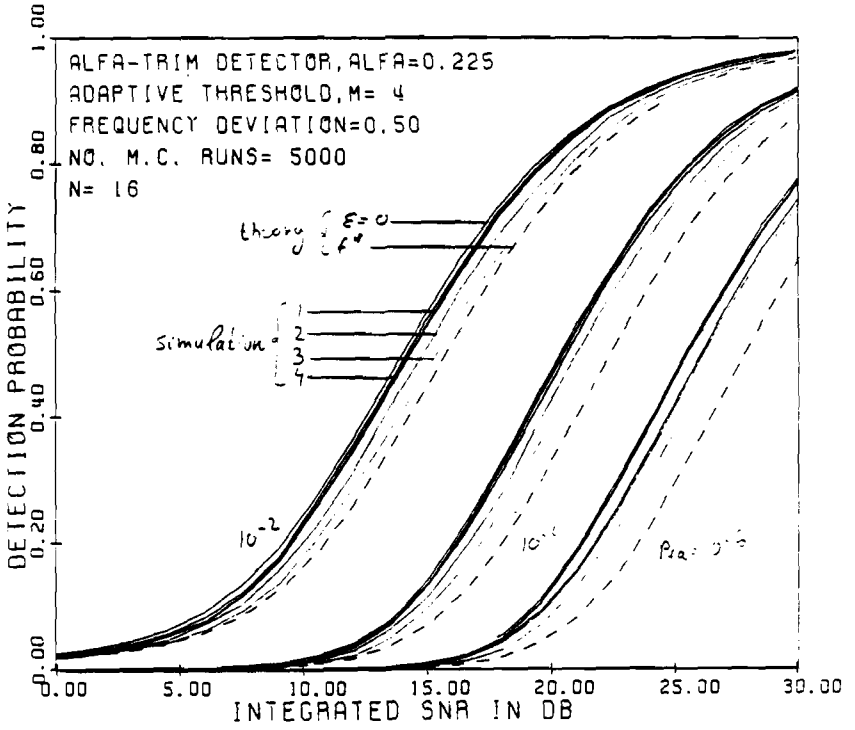


Fig. 5.26 Detection probability. Test - D2, α=.225,Δf =.5

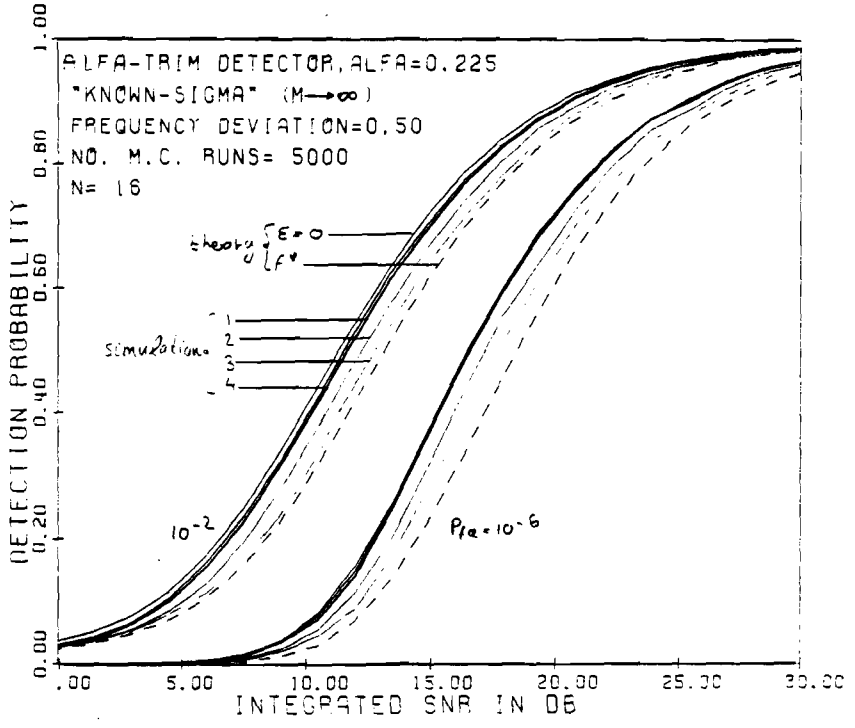
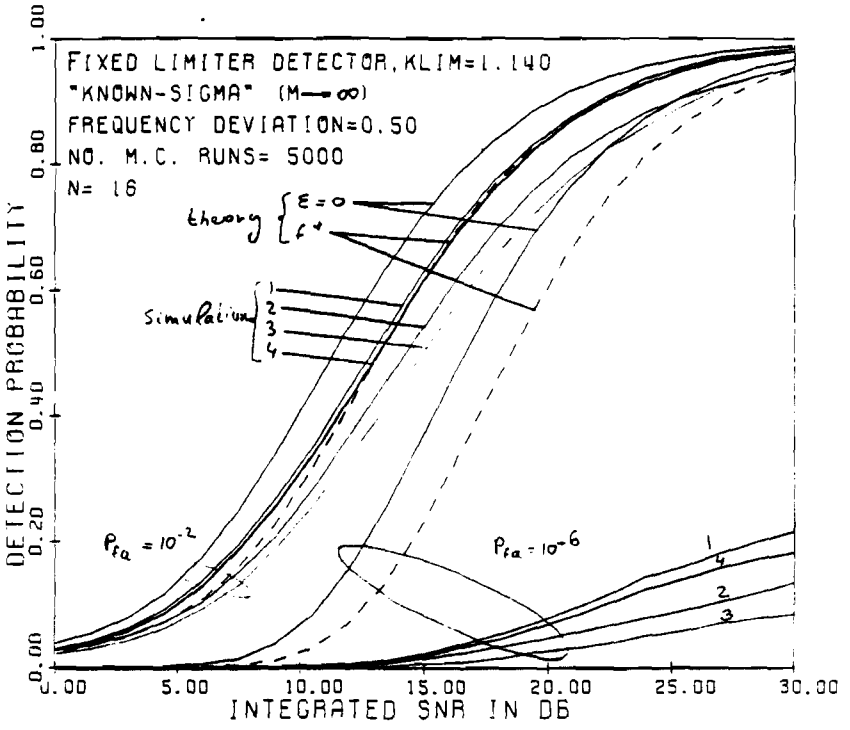


Fig. 5.27 Detection probability. Test - D5, K=1.14,Δf =.5



Appendix A

Small Sample Performance of the Sign Detector

The sign detector (SD) is the simplest and probably the most popular non-parametric detector for a deterministic lowpass signal in noise with uncertain symmetrical density. Under several assumptions, it is known to be the UMP detector [32] or the minimax LORD detector [5] for those particular situations. Under the asymptotic weak-signal/large-sample ($n \rightarrow \infty, H_1 \rightarrow H_0$) assumptions, the ARE with respect to the linear detector (LD) at noise density $f(\cdot)$ is $ARE_{SD,LD} \approx 4(\sigma f(0))^2$. This is 0.637 for Gaussian noise, 2 for Laplace (double exponential) noise and in between for many other densities. Hence, the SD is generally considered to be very efficient for many situations of uncertainty.

Small sample analysis reveals that it may not perform as well, as the fine details of the actual small sample-size error probabilities are lost in the asymptotic analysis. The sign test is given by:

$$T(\mathbf{x}) = \sum_{i=1}^{n} u(x_i) \begin{cases} > t & , \quad H_1 \\ = t & , \quad H_1 \text{ with probability } c \\ < t & , \quad H_0 \end{cases} \tag{A.1}$$

where $u(\cdot)$ is the unity step function. Notice that a randomized test is required since $T(\mathbf{x})$ is a discrete binomially distributed r.v. Assuming symmetry of $f(\cdot)$, the false alarm probability is given by:

$$\alpha_0 = (\frac{1}{2})^n \begin{cases} \sum_{k=t+1}^{n} \begin{pmatrix} n \\ k \end{pmatrix} + c \begin{pmatrix} n \\ t \end{pmatrix} & , \quad t \leq n-1 \\ c & , \quad t = n \end{cases} \tag{A.2}$$

When n is small and the desired α_0 is very small, the second row of (A.2) must be chosen, and therefore the detection probability is also proportional to the randomization constant c which might be very small. Denoting $Prob_{H_1}\{x_i \geq 0\} = p$, it is given by

$$P_d = cp^n = \alpha_0 2^n p^n \qquad \leq \qquad \alpha_0 2^n \mid_{SNR \rightarrow \infty} \tag{A.3}$$

The critical values of the sample size such that for $n \leq n_c$ the detector becomes very poor ($P_d \leq 0.5$, *even* for $SNR \rightarrow \infty$) are found to be 12, 19 and 26 for $\alpha_0 = 10^{-4}, 10^{-6}, 10^{-8}$, respectively; these numbers are typical of many practical systems. A similar situation is encountered with the Wilcoxon detector, as its minimal false alarm probability (without randomization) is also 2^{-n}.

Appendix B

An Approximation for the Distribution of the Envelope Detector

In this appendix an approximation is derived for the distribution of the coherent envelope statistic: $R = I^2 + Q^2$, $I = (1/n)\sum_1^n x_i$, $Q = (1/n)\sum_1^n y_i$, where $\mathbf{x}$ and $\mathbf{y}$ are i.i.d. vectors which are also independent of each other. When the samples are Gaussian, R is one sided exponentially distributed. Thus, we are interested in an approximation which exhibits this exponential limiting behavior.

Formally, any univariate density $f(x)$ can be expanded in a series of orthonormal functions $\{\Phi_k(x)\}$ as

$$f(x) = \sum_{k=0}^{\infty} C_k \, \Phi_k(x) \tag{B.1}$$

which is L_2 convergent when all moments of f exist. A classical expansion in terms of the Gaussian density and its derivatives (which are proportional to Hermite polynomials) is the type A Gram-Charlier series [36], or its reordered version, the Edgeworth's form which, when applied to the distribution of the samples mean, gives the best order in powers of $n^{-1/2}$. Conditions for convergence of the infinite Edgeworth series are rather technical and implicit (they require conditions on f which is usually unknown); moreover, they are of little value from the statistical viewpoint where the important question is whether a *small* number of terms can furnish a reasonable approximation. In particular, while the Edgeworth expansion of the sample mean converges to the Gaussian density with finite terms as $n \to \infty$, when the expansion is carried for finite n, adding more terms in the series above some optimal number usually results in a *worse fit*. See [36,Chap. 6] and [37,Chap. 16].

As pointed out in a footnote [36,section 6.24], when an approximation for the distri-

bution of a *positive* random variable is desired, it is plausible to use a family of functions which are supported only on $[0, \infty)$. In this way, the error from the tails of the Gaussian density and its derivatives at $x < 0$ would be eliminated. Since in our case the limiting one sided exponential distribution belongs to the Gamma family, the appropriate polynomials are those of Laguerre defined by [38]

$$L_i^a(x) = \frac{e^x x^{-a}}{i!} \frac{d^i}{dx^i}(e^{-x} x^{i+a}) \tag{B.2}$$

The orthonormality relation is

$$\int_0^\infty e^{-x} x^a L_i^a(x) L_j^a(x) = \frac{\Gamma(a+1+i)}{i!} \delta_{ij} \tag{B.3}$$

hence the expansion is

$$f(x) = \sum_{i=0}^{\infty} C_i \; e^{-x} x^a L_i^a(x) \tag{B.4}$$

where

$$C_i = \frac{i!}{\Gamma(a+i+1)} \int_0^\infty L_i^a(x) f(x) dx \tag{B.5}$$

The first four polynomials are

$$L_0^a(x) = 1 \tag{B.6}$$

$$L_1^a(x) = 1 + a - x \tag{B.7}$$

$$L_2^a(x) = \frac{1}{2}[(a+1)(a+2) - 2x(a+2) + x^2] \tag{B.8}$$

$$L_3^a(x) = \frac{1}{6}[(a+1)(a+2)(a+3) - 3x(a+2)(a+3) + 3x^2(a+3) - x^3] \tag{B.9}$$

Like the Edgeworth expansion, it is possible to eliminate the terms $i = 1, 2$ by a transformation. Defining $x = by$, the following expressions are obtained from (B.5) and (B.6)-(B.9):

$$C_0 = \frac{1}{b\,\Gamma(a+1)} \tag{B.10}$$

$$C_1 = \frac{1}{b\,\Gamma(a+2)}\left[(a+1) - \frac{\mu_1}{b}\right] \tag{B.11}$$

$$C_2 = \frac{1}{b\,\Gamma(a+3)}\left[(a+2)(a+1) - 2(a+2)\frac{\mu_1}{b} + \frac{\mu_2}{b^2}\right] \tag{B.12}$$

$$C_3 = \frac{1}{b\,\Gamma(a+4)}\left[(a+3)(a+2)(a+1) - 3(a+3)(a+2)\frac{\mu_1}{b^2} + 3(a+3)\frac{\mu_2}{b^2} - \frac{\mu_3}{b^3}\right] \tag{B.13}$$

where $\mu_i = E_x(x^i)$. Solving for $C_1 = C_2 = 0$, we obtain:

$$b = \frac{\mu_2 - \mu_1^2}{\mu_1} \triangleq \frac{1}{\beta} \tag{B.14}$$

$$a = \mu_1^2/(\mu_2 - \mu_1^2) - 1 \triangleq \alpha \tag{B.15}$$

Hence, upon substituting back with $y = \beta x$

$$f(x) = \frac{\beta^\alpha e^{-\beta x} x^{\alpha-1}}{\Gamma(\alpha)} + \frac{\beta^2}{\Gamma(\alpha+3)}\left[(\alpha+2)\mu_2 - \beta\mu_3\right]\beta^\alpha e^{-\beta x} x^{\alpha-1} L_3^{\alpha-1}(\beta x) + \cdots \tag{B.16}$$

The leading term is recognized as the Gamma density which, as $\alpha \to 1$, tends to the desired exponential density.

Integrating (B.16) to get the false alarm probability $P_r\{x > t\,\mu_1\} = P_r\{y > t\,\alpha\}$ is simplified by using the defining expression for the Laguerre polynomials (B.2). Some tedious algebra yields:

$$P_{fa} = P_r\{x > t\,\mu_1\} = [1 - I(\alpha, t\,\alpha)] + \frac{C_3'}{6}(\alpha, \mu_1, \mu_3) e^{-\alpha t}(\alpha t)^\alpha \left[2\alpha t(\alpha+2) - (\alpha t)^2 - (\alpha+1)(\alpha+2)\right] + \cdots \tag{B.17}$$

where

$$C_3' = \frac{C_3}{\beta} = \frac{\alpha}{\Gamma(\alpha+3)}\left[(\alpha+1)(\alpha+2) - \frac{\alpha^2\mu_3}{\mu_1^3}\right] \tag{B.18}$$

and

$$I(\alpha, \alpha t) \triangleq \int_0^{\alpha t} \frac{e^{-t} t^{\alpha-1} dt}{\Gamma(\alpha)} \tag{B.19}$$

is the incomplete Gamma function. It is clear from (B.17) that the second term vanishes for $t = 0$ as well as for $t = \infty$, thus fulfilling some of the properties of a proper distribution (since the first term is a proper distribution). However, it cannot be said whether or

not (B.16) can assume negative values, a common feature of all finite Gram-Charlier expansions. The following term in the expansion turns to be so complicated so it will not be given here.

To get specific results for the coherent envelope statistics R defined in the beginning of the appendix, we need to express the moments of R in terms of those of the i.i.d. samples, $E(R^m) = \frac{1}{n^m} E[(\sum_1^n x_i)^2 + (\sum_1^n y_i)^2]^m$. Using the moment generating function, tedious but straightforward manipulations yield

$$\mu_1 = E(R) = \frac{2}{n}\sigma_i^2 \tag{B.20}$$

$$\mu_2 = E(R^2) = \frac{2\sigma_i^4}{n^2}[4 + \frac{K}{n}] \quad , \quad K \triangleq \frac{m_4}{\sigma_i^4} - 3 \tag{B.21}$$

$$\mu_3 = E(R^3) = \frac{2\sigma_i^6}{n^3}[24 + \frac{18K}{n} + \frac{1}{n^2}(L - 15K)] \quad , \quad L \triangleq \frac{m_6}{\sigma_i^6} - 15 \tag{B.22}$$

where $\sigma_i^2 = E(x_i^2)$, $m_j = E(x_i^j)$ and K is the coefficient of kurtosis. When x_i are Gaussian, both K and L obviously vanish. Substituting these into (B.15) and (B.18) we obtain:

$$\alpha = (1 + K/2n)^{-1} \tag{B.23}$$

$$C_3' = \frac{\alpha}{\Gamma(\alpha + 3)}[3\alpha + 2 - \alpha^2(5 + \frac{9K}{2n} + \frac{L - 15K}{4n^2})] \tag{B.24}$$

Hence, the limiting behavior of the approximation with only the first four terms of (B.4) occurs when the number of samples n is much greater than the kurtosis ($K/n \rightarrow 0$). Then $\alpha \rightarrow 1$ and the first term in (B.17) approaches e^{-t} as required by the central limit theorem, though (unlike the Edgeworth expansion) the first term itself actually includes all powers of K/n in a highly nonlinear fashion. When $\alpha < 1$, a comparison with tables of the incomplete Gamma function shows that for sufficiently large t (such that the nominal false alarm probability $P_{fa} = e^{-t}$ is small), the outcome of Eq. (B.17) could be orders of magnitude higher than the P_{fa} as expected. In the second term of (B.17), all the expressions tend to a constant when $\alpha \rightarrow 1$ except C_3' which

vanishes identically when $\alpha = 1$. For other values of α, expansion in K/n gives

$$\lim_{\frac{K}{n} \to 0} C_3' = -\frac{K}{n} \frac{11}{12} + 0(\frac{K^2}{n^2}) \tag{B.25}$$

For the normal-normal mixture: $f(x_i) = (1 - \epsilon) N(0, 1) + \epsilon N(0, C^2)$,

$$K = 3\epsilon(1 - \epsilon)\left[\frac{C^2 - 1}{1 - \epsilon + \epsilon C^2}\right]^2 \underset{\epsilon C^2 >> 1}{\longrightarrow} \frac{3}{\epsilon} \tag{B.26}$$

$$L = 15\left[\frac{(1 - \epsilon + \epsilon C^6)}{(1 - \epsilon + \epsilon C^2)^3} - 1\right] \underset{\epsilon C^2 >> 1}{\longrightarrow} \frac{15}{\epsilon^2} \tag{B.27}$$

Therefore, with small amounts of large contamination and n not very large, α can be much smaller than one and the contribution of (B.24) is also not negligible.

Actual computation of (B.17), using equations (B.20)-(B.24), (B.26), (B.27), is shown in figures 2.1-2.5 of section 2.2 and compared to Monte-Carlo simulation results. For α roughly in the range [0.5,1], this approximation is seen to be quite good. It is still a reasonable prediction of the amount of increase in the false alarm probability down to $\alpha = 0.3$.

APPENDIX C

The Maximin Test for a Single Observation

For a single observation of a r.v. R which comes from simple mixture hypotheses $\mathbf{P}_0$ and $\mathbf{P}_1$ as in Eq. (1.1), Huber's [2] maximin robust test takes the form

$$d^*(R) = \begin{cases} H_1 & , \quad l(R\,;L'\,,L''\,) > t \\ H_1 \text{ with probability } c & , \quad l(R\,;L'\,,L''\,) = t \\ H_0 & , \quad l(R\,;L'\,,L''\,) < t \end{cases} \tag{C.1}$$

where the soft limited $l(R\,;L'\,,L''\,)$ is given by Eqs. (1.6-1.8). The constants $0 \leq L' \leq L'' < \infty$ are found by solving

$$P_1\{L \geq L'\} + L'\,P_0\{L < L'\} = (1-\epsilon)^{-1} \tag{C.2}$$

$$P_0\{L < L''\} + (1/L'')P_1\{L \geq L''\} = (1-\epsilon)^{-1} \tag{C.3}$$

where $L(R) = f_1(R)/f_0(R)$ is the LR under the nominal situation, and is assumed to be continuous. Application of these equations to the distribution of the coherent envelope of a Rayleigh narrowband signal in narrowband Gaussian noise, Eqs. (3.3-3.4), results in Eqs. (3.6-3.7).

Since $l(R\,;L'\,,L''\,)$ is a monotone increasing function of $L(R)$, which in turn is monotonic in R, Eq. (C.1) can be reformulated as a randomized test on R. The threshold and the randomization constant are functions of the least favorable p.d.f.'s. Three different cases are possible, according to $t = L'$, $L' < t < L''$ and $t = L''$, see Figure 1. Define R' and R'' by $L(R') = L'$, $L(R'') = L''$. Define also $\alpha(x) = \int_x^\infty f_0(r)dr$, $\beta(x) = \int_x^\infty f_1(r)dr$. Then,

$$\underline{\text{Case A}} \qquad t = L' \;\Rightarrow\; \epsilon < (\alpha_0 - \alpha(R'))/(1 - \alpha(R')) \tag{C.4}$$

$$d^*(R) = \begin{cases} H_1 & , \quad R \geq R' \\ H_1 \text{ with probability } c & , \quad R < R' \end{cases} \tag{C.5}$$

where $0 \leq c \leq 1$ is given, as a function of the desired α_0 and of R' which solves (C.2), by

$$c = \frac{\alpha_0 - \epsilon - (1-\epsilon)\alpha(R')}{(1-\epsilon)(1-\alpha(R'))} \tag{C.6}$$

The detection probability under the least-favorable q_1^* is

$$\beta(d^*, q_1^*) = c + (1-\epsilon)(1-c)\,\beta(R') = 1 - (1-\alpha_0)L' \tag{C.7}$$

Case B $$L < t < L'' \Rightarrow \frac{\alpha_0 - \alpha(R')}{1-\alpha(R')} \leq \epsilon < \frac{\alpha_0 - \alpha(R'')}{1-\alpha(R'')} \tag{C.8}$$

$$d^*(R) = \begin{cases} H_1\,, & R \geq R_t \\ H_0\,, & R < R_t \end{cases} \tag{C.9}$$

where

$$\alpha_0 = \epsilon + (1-\epsilon)\,\alpha(R_t) \tag{C.10}$$

Here,

$$\beta(d^*, q_1^*) = (1-\epsilon)\,\beta(R_t) \tag{C.11}$$

Case C $$L'' = t \Rightarrow \epsilon \geq (\alpha_0 - \alpha(R''))/(1-\alpha(R'')) \tag{C.12}$$

$$d^*(R) = \begin{cases} H_1 \text{ with probability } c\,, & R \geq R'' \\ H_0\,, & R < R'' \end{cases} \tag{C.13}$$

where

$$c = \frac{\alpha_0}{\epsilon + (1-\epsilon)\,\alpha(R'')} \tag{C.14}$$

Here,

$$\beta(d^*, q_1^*) = \frac{\alpha_0(1-\epsilon)\,\beta(R'')}{\epsilon + (1-\epsilon)\,\alpha(R'')} = \alpha_0 L'' \tag{C.15}$$

Notice that in this case where large ϵ necessitates randomized maximin test, the resulting maximin lower bound on β, Eq. (C.15) is very small, as it is proportional to the desired α_0.

The proof of Eqs. (C.4-C.15) follows from (C.1-C.3). For Case A, (C.1) is translated into (C.5) as a result of the above mentioned monotonicity properties. To satisfy the constraint on the probability of false alarm, we must have

$$\alpha_0 = P_0^*\{R \geq R'\} + c\, P_0^*\{R < R'\}$$
$$= P_0^*\{R' \leq R < R''\} + P_0^*\{R \geq R''\} + cP_0^*\{R < R'\}$$

Utilizing the expression for the least favorable density q_0^*, Eq. (1.6),

$$\alpha_0 = (1-\epsilon)P_0\{R' \leq R \leq R''\} + c\,(1-\epsilon)P_0\{R < R'\} + \frac{(1-\epsilon)}{L''} P_1\{R \geq R''\}$$

The last term is equal to $1 - (1-\epsilon)P_0\{R < R''\}$, by virtue of Eq. (C.3). Collecting terms, we obtain

$$\alpha_0(1-\epsilon)(1-\alpha(R'))(c-1) + 1$$

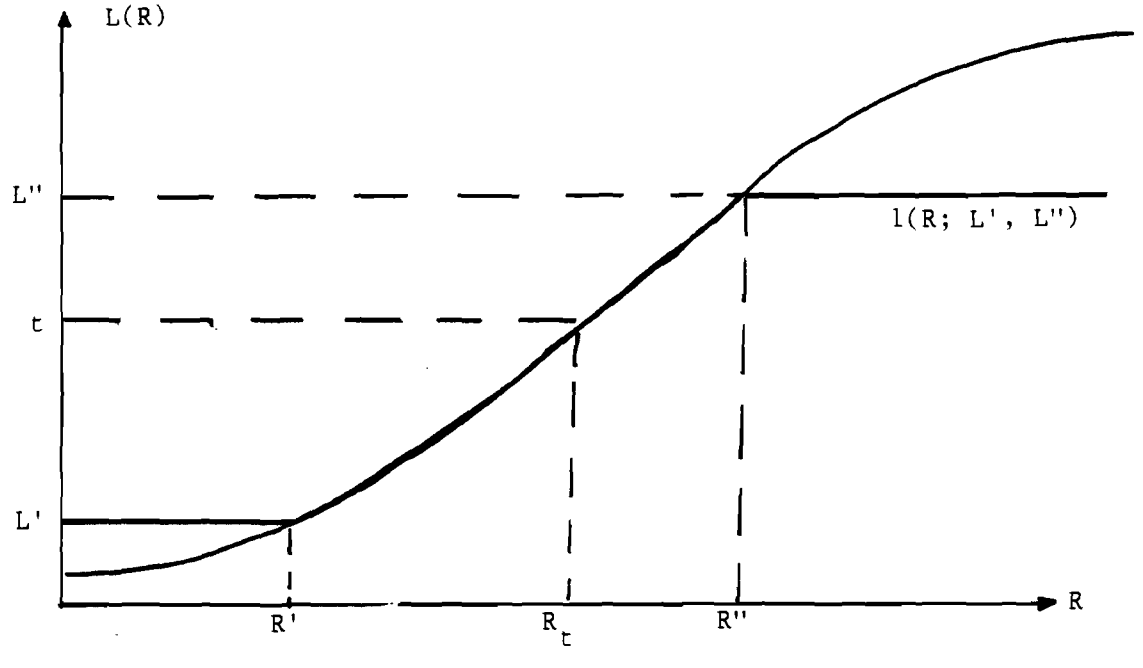
from which (C.6) follows. In order to satisfy $c \geq 0$, we get the appropriate ϵ range, Eq. (C.4). In a similar manner, from Eq. (1.2) for q_1^*,

$$\beta(d^*, q_1^*) = P_1^*\{R \geq R'\} + cP_1^*\{R < R'\}$$
$$= (1-\epsilon)P_1\{R \geq R'\} + (1-\epsilon)\, L' c P_0\{R < R'\}$$

and the last term, by virtue of Eq. (C.2), is equal to $c\,[1 - (1-\epsilon)P_1\{R \geq R'\}]$. Collecting terms, the first equality of Eq. (C.7) is obtained, and the second one is found by substitution of (C.6) and using (C.2) again.

Cases B and C are proved in a similar manner.

Figure 1



Appendix D

Statistics of the SSQME Test with Unmatched Frequency

We first generalize the asymptotic normality of M-estimators for a sequence of r.v.'s which are not identically distributed.

Lemma

Let $l \in \Psi$ of Section 4.1 and assume also that l is bounded. Let $\{x_i\}$ be a sequence of independent r.v.'s with p.d.f. $\frac{1}{\sigma_i} f\left(\frac{x_i - Am_i}{\sigma_i}\right)$ where f is symmetric and absolutely continuous, and $\lim_{n\to\infty} nA^2 = a^2 < \infty$. Then $\sqrt{n}\,\hat{A}_n$, where $\hat{A}_n = arg\left\{\sum_{i=1}^{n} l(x_i - \hat{A}_n) = 0\right\}$ is asymptotically distributed as $N(m, V)$ where

$$m = \frac{a\sum m_i E_i(l')}{\sum E_i(l')} \quad , \quad E_i(l') \triangleq \int l'(x) f\left(\frac{x}{\sigma_i}\right)\frac{dx}{\sigma_i} \tag{D.1}$$

and

$$V = \frac{\frac{1}{n}\sum E_i(l^2)}{\left[\frac{1}{n}\sum E_i(l')\right]^2} \quad , \quad E_i(l^2) \triangleq \int l^2(x) f\left(\frac{x}{\sigma_i}\right)\frac{dx}{\sigma_i} \tag{D.2}$$

Proof The proof follows Huber's [1] with some modifications to account for the unequal means and variances, hence it will only be briefly sketched. By monotonicity of $l(\cdot)$,

$$\lim_{n\to\infty} Prob\left\{\hat{A}_n \sqrt{n} \le k\right\} = \lim_{n\to\infty} Prob\left\{\sum l(x_i - kn^{-1/2}) \le 0\right\} \tag{D.3}$$

Define $y_{n_i} = l(x_i - kn^{-1/2})$. Upon expanding in a Taylor series of powers of $n^{-1/2}$, and invoking the assumed symmetry properties of f and l, some terms vanish and one obtaines:

$$m_{n_i} \triangleq E(y_{n_i}) = n^{-1/2}(am_i - k)E_i(l') + O(n^{-1}) \tag{D.4}$$

$$v_{n_i} \triangleq Var(y_{n_i}) = E_i(l^2) + O(n^{-1}) \tag{D.5}$$

There is a slight complication since the y_{n_i} are different for different values of n, and the

usual formulations of normal convergence do not apply. However, this case is covered by the extended Lindeberg - Feller theorem [41], which asserts that a necessary and sufficient condition for convergence to a standard normal distribution of $s_n = \sum_{i=1}^{r_n} (y_{n_i} - m_{n_i})/\sigma_n$ with $\sigma_n^2 = \sum_{i=1}^{r_n} v_{n_i}$, is that each summand contributes negligibly to s_n in the sense of Lindeberg condition

$$\lim_{n \to \infty} \max_{1 \le i \le r_n} \frac{v_{n_i}}{\sigma_n^2} = 0 \tag{D.6}$$

Since $l(\cdot)$ is bounded and $v_{n_i} \neq 0$, σ_n^2 grows as n and the condition is satisfied. Interchanging of $\lim_{n \to \infty}$ and $Prob\{\cdot\}$ is allowed since $\{\hat{A}_n\}$ converges. Subtracting $\lim_{n \to \infty} \sum m_{n_i}/\sigma_n = (m - k)/V$ from both sides of the argument of (D.3) yields

$$\lim_{n \to \infty} Prob\{\hat{A}_n \sqrt{n} \le k\} = \Phi\left(\frac{k - m}{V}\right) \tag{D.7}$$

We now apply the lemma to the case of frequency deviation treated in section 5.3. The means of the quadrature noise components are given by $E(x_i) = A\cos(\phi + \phi_i)$ and $E(y_i) = A\sin(\phi + \phi_i)$ where $\phi_i = i\,2\pi f_d / f_r$, and the variances are identical. Denote the M-estimators by I and Q for the x and y channels, respectively. They are jointly Gaussian with equal variance $V = E(l^2)/E^2(l')$ (since the noise components have identical variance), and their means are $E(I) = \sqrt{n} A \sum\cos(\phi + \phi_i)$ and $E(Q) = \sqrt{n} A \sum\sin(\phi + \phi_i)$ from (D.1). The argument of the exponent of the joint Gaussian density is thus

$$\frac{-1}{2V}\{ I^2 + Q^2 + E^2(I) + E^2(Q) - 2[E(I)I + E(Q)Q] \}$$

The middle pair does not depend on ϕ

$$\frac{E^2(I) + E^2(Q)}{nA^2} = [\sum\cos(\phi + \phi_i)]^2 + [\sum\sin(\phi + \phi_i)]^2 = \left|\sum e^{j(\phi + \phi_i)}\right|^2 = \cdots$$

$$\cdots = [\sum c_i]^2 + [\sum s_i]^2 = \frac{\sin^2(n\pi f_d / f_r)}{\sin^2(\pi f_d / f_r)} \triangleq G(f_d)$$

where $c_i \triangleq \cos(\phi_i)$ and $s_i \triangleq \sin(\phi_i)$, by straightforward calculation. Also,

$$\frac{IE(I) + QE(Q)}{\sqrt{n}A} = \sqrt{C^2+S^2}\cos(\phi-\tan^{-1}(\frac{S}{C}))$$

where $C \triangleq I\sum c_i + Q\sum s_i$ and $S \triangleq -I\sum s_i + Q\sum c_i$, from which $C^2+S^2 = (I^2+Q^2)G(f_d)$ follows as can be verified by substituting for the definitions. We finally have

$$f(I,Q \mid A,\phi) = \frac{1}{2\pi V}\exp\left[\frac{I^2+Q^2+nA^2G(f_d)-2\sqrt{n}A\sqrt{(I^2+Q^2)G(f_d)}\cos(\phi-\tan^{-1}(\frac{S}{C}))}{-2V}\right] \tag{D.8}$$

Averaging this with respect to ϕ on $[0,2\pi]$ gives the same expression as in section 4.2, since S and C do not depend on ϕ, and where the effective SNR is replaced by $nA^2G(f_d)/2V$. Hence, the detection probabilities for the various SSQME tests of Section 5.2 are identical with the case $f_d = 0$ if the effective SNR is attenuated by $G(f_d)$.

In the FFT processor, the sidelobes can be reduced, and the response can be flattened over the mainlobes of the individual filters ($\mid nf_d/f_r \mid \leq 0.5$), at a price of reduced SNR at $f_d = 0$, by appropriate weighting (e.g., Hamming's):

$\{\mathbf{x},\mathbf{y}\} \rightarrow \{\mathbf{w}^T\mathbf{x},\mathbf{w}^T\mathbf{y}\}$, which changes $G(f_d) \rightarrow \mid \sum_{k=1}^{n} w_k \exp(j2\pi k f_d/f_r) \mid^2$. The same can be done with the SSQME test. The derivation of the test statistic is similar to that above, but now we have to account for the unequal variances as a consequence of the weighting. The frequency response will be given by $G(f_d) = \mid \sum_k w_k' \exp(j2\pi k f_d/f_r) \mid^2$, where $w_k' \triangleq E_k(l')/\sum_k E_k(l')$. The design of the weighting coefficients is more complicated than in the linear FFT case; first $\{w_k'\}$ are chosen for the desired response, and then $\int l'(w_k x) f(x) dx = w_k'$ has to be solved for $\{w_k\}$.

REFERENCES

1. P.J. Huber, "Robust estimation of a location parameter," *Ann. Math. Stat.*, vol. 35, pp. 73-101, 1964.

2. P.J. Huber, "A robust version of the probability ratio test," *Ann. Math. Stat.*, vol. 36, pp. 1753-1758, 1965.

3. P.J. Huber, "Robust Statistics," Wiley & Sons, New York, 1981.

4. R.D. Martin and S.C. Schwartz, "Robust detection of a known signal in nearly Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 50-56, January 1971.

5. S.A. Kassam and J.B. Thomas, "Asymptotically robust detection of a known signal in contaminated non-Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 22-26, January 1976.

6. A.H. El-Sawy and V.D. VandeLinde, "Robust detection of known signals," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 722-727, November 1977.

7. H.V. Poor and J.B. Thomas, "Asymptotically robust quantization for detection," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 222-229, 1978

8. S.A. Kassam, "Locally robust array detectors for random signals," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 309-316, May 1978.

9. H.V. Poor, M. Mami and J.B. Thomas, "On robust detection of discrete-time stochastic signals," *J. Franklin Inst.*, vol. 309, pp. 29-53, January 1980.

10. S.A. Kassam, G. Moustakides and J.G. Shin, "Robust detection of known signals in asymmetric noise," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 84-91, 1982.

11. J.G. Shin and S.A. Kassam, "Robust detector for narrowband signals in non-Gaussian noise," *J. Acoust. Soc. Amer.*, vol. 74, pp. 527-533, August 1983.

12. S.A. Kassam, "Detection of narrowband signals: Asymptotic robustness and quantization," *1984 Conference on Information Sciences and Systems*, Princeton University, March 1984.

13. J.V. DiFranco and W.L. Rubin, "Radar detection," Prentice Hall, New Jersey, 1968.

14. H.L. VanTrees, "Detection, estimation and modulation theory; Part III: Radar-Sonar signal processing and Gaussian signals in noise," Wiley & Sons, New York, 1971.

15. C.W. Helstrom, "Statistical theory of signal detection," Second Edition, Pergamon Press, Oxford, 1968.

16. J.B.G. Roberts, "The impact of new VLSI technology on radar signal processing," *Proc. of the International Conference on Radar,* pp. 9-15, Paris, 1984.

17. P. Swerling, "Probability of detection for fluctuating targets," Rand Report RM-1217, March 1954. (Reprinted in *IRE Trans. Inform. Theory,* vol. IT-6, pp. 269-308, April 1960.)

18. H.M. Finn and R.S. Johnson, "Adaptive detection mode with threshold control as function of spatially sampled clutter-level estimates," *RCA Review,* vol. 30, pp. 414-465, September 1968.

19. V.G. Hansen, "Constant false alarm rate processing in search radars," *Proc. of the IEEE 1973 Inter. Radar Conf.*, pp. 325-332, London, 1973.

20. M. Weiss, "Analysis of some modified cell-averaging CFAR processors in multiple target situations," *IEEE Trans. on Aeros. Elect. Syst.*, vol. AES-18, no. 1, pp. 102-114, January 1982.

21. J.D. Gibson and J.L. Melsa, "Introduction to nonparametric detection with applications," Academic Press, New York, 1975.

22. A.A. Winder, "Underwater sound - a review; Part II: Sonar system technology," *IEEE Trans. on Sonics and Ultrasonics,* vol. SU-22, no. 5, pp. 291-332, 1975.

23. A. Wald, "Asymptotically most powerful tests of statistical hypotheses," *Ann. Math. Statist.*, vol. 12, pp. 1-19, March 1941.

24. H. Chernoff, "On the distribution of the likelihood ratio," *Ann. Math. Statist.*, vol. 25, no. 3, pp. 573-578, September 1954.

25. J. Capon, "On the asymptotic relative efficiency of locally optimum detectors," *IRE Trans. Inform. Theory,* vol. IT-7, pp. 67-72, April 1961.

26. D. Middleton, "An introductionto statistical communication theory," McGraw-Hill, New York, 1960.

27. D.A.S. Fraser, "Nonparametric methods in statistics," Academic Press, New York, 1951.

28. T.S. Ferguson, "Mathematical statistics," Academic Press, New York, 1967.

29. A.B. Martinez, "Asymptotic performance of detectors with non-zero input SNR," *Proc. 20th Annual Allerton Conf. on Commun., Control and Computing,* University of Illinois, Urbana-Champaign, IL, pp. 749-758, 1982.

30. D.L. Michalsky, G.L. Wise and H.V. Poor, "A relative efficiency study of some popular detectors," *J. of Franklin Inst.*, vol. 313, pp. 135-148, March 1982.

31. A.D. Spaulding, "Locally optimum and suboptimum detector performance in non-Gaussian noise," *Proc. of IEEE ICC 1982,* vol. 7, pp. 2H.2.1-7, June 1982.

32. J.B. Thomas, "Nonparametric detection," *Proc. IEEE,* vol. 58, no. 5, pp. 623-631, 1970.

33. P.J. Bickel, "On some robust estimates of location," *Ann. Math. Statist.,* 36, pp. 847-858, 1965.

34. F.R. Hampel, "The influence curve and its role in robust estimation," *J. Amer. Statist. Ass.,* 62, pp. 1179-1186, 1974.

35. D.F. Andrews, et.al., "Robust estimation of location, survey and advances," Princeton, New Jersey, Princeton University Press, 1972.

36. S.M. Kendal and A. Stuart, "The advanced theory of statistics," vol. 1, Fourth Edition, Charles Griffin & Company, London, 1977.

37. W. Feller, "An introduction to probability theory and its applications," vol. II, Wiley & Sons, New York, 1966.

38. G. Szego, "Orthogonal Polynomials," New York, 1939.

39. A.V. Aho, J.E. Hopcroft and J.D. Ullman, "The design and analysis of computer algorithms," Addison-Wesley, Reading, MA, 1974.

40. V.G. Hansen, "Detection performance of the narrow-band Wilcoxon detector against Gaussian noise," *IEEE Trans. on Inform. Theory,* vol. IT-18, pp. 664-667, September 1972.

41. P. Billingsley, "Probability and measure," Wiley & Sons, New York, 1979.

42. C.A. Field and F.R. Hampel; "Small-sample asymptotic distribution of M-estimates of location," *Biometrika,* vol. 69, no. 1, pp. 29-46, 1982.

43. H.E. Daniels, "Saddlepoint approximations in statistics," *Ann. Math. Statist.,* vol. 25, pp. 631-650, 1954.

44. R.L. Mitchell, "Importance sampling applied to simulation of false alarm statistics," *IEEE Trans. on Aeros. Elect. Syst.,* vol. AES-17, no. 1, pp. 15-24, January 1981.

45. M. Weiss and S.C. Schwartz, "On optimal minimax jamming and detection of radar signals," *accepted for publication, IEEE Transactions on Aerospace and Electronic Systems.*

46. L.A. Jaeckel, "Robust estimates of location: Symmetric and asymmetric contamination", *Ann. Math. Statist.,* 42, pp. 1020-1034, 1971.

47. S.M. Stigler, "Linear functions of order statistics with smooth weight functions", *Ann. Statistics,* 2, pp. 676-693, 1974.

48. M. Weiss and S.C. Schwartz, "Robust detection of coherent radar signals in nearly Gaussian noise", Proc. of the IEEE International Radar Conference, Washington,

May 1985.

49 J. Hajek and Z. Sidak, "Theory of rank tests", Academic Press, New York, 1967.

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

FEBRUARY 1982

	Copies
Statistics and Probability Program (Code 411(SP)) Office of Naval Research Arlington, VA 22217	3
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12
Commanding Officer Office of Naval Research Eastern/Central Regional Office Attn: Director for Science Barnes Building 495 Summer Street Boston, MA 02210	1
Commanding Officer Office of Naval Research Western Regional Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1
U. S. ONR Liaison Office - Far East Attn: Scientific Director APO San Francisco 96503	1
Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1
Commandant of the Marine Coprs (Code AX) Attn: Dr. A. L. Slafkosky Scientific Advisor Washington, DC 20380	1
Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1
U. S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1
Director National Security Agency Attn: R51, Dr. Maar Fort Meade, MD 20755	1
ATAA-SL, Library U.S. Army TRADOC Systems Analysis Activity Department of the Army White Sands Missile Range, NM 88002	1
ARI Field Unit-USAREUR Attn: Library c/o ODCSPER HQ USAEREUR & 7th Army APO New York 09403	1
Library, Code 1424 Naval Postgraduate School Monterey, CA 93940	1
Technical Information Division Naval Research Laboratory Washington, DC 20375	1
OASD (I&L), Pentagon Attn: Mr. Charles S. Smith Washington, DC 20301	1

Copies

Director
AMSAA
Attn: DRXSY-MP, H. Cohen
Aberdeen Proving Ground, MD 21005 1

Dr. Gerhard Heiche
Naval Air Systems Command (NAIR 03)
Jefferson Plaza No. 1
Arlington, VA 20360 1

Dr. Barbara Bailar
Associate Director, Statistical Standards
Bureau of Census
Washington, DC 20233 1

Leon Slavin
Naval Sea Systems Command (NSEA 05H)
Crystal Mall #4, Rm. 129
Washington, DC 20036 1

B. E. Clark
RR #2, Box 647-B
Graham, NC 27253 1

Naval Underwater Systems Center
Attn: Dr. Derrill J. Bordelon
Code 601
Newport, Rhode Island 02840 1

Naval Coastal Systems Center
Code 741
Attn: Mr. C. M. Bennett
Panama City, FL 32401 1

Naval Electronic Systems Command (NELEX 612)
Attn: John Schuster
National Center No. 1
Arlington, VA 20360 1

Defense Logistics Studies Information Exchange
Army Logistics Management Center
Attn: Mr. J. Dowling
Fort Lee, VA 23801 1

Copies

Reliability Analysis Center (RAC)
RADC/RBRAC
Attn: I. L. Krulac
Data Coordinator/
Government Programs
Griffiss AFB, New York 13441 1

Technical Library
Naval Ordnance Station
Indian Head, MD 20640 1

Library
Naval Ocean Systems Center
San Diego, CA 92152 1

Technical Library
Bureau of Naval Personnel
Department of the Navy
Washington, DC 20370 1

Mr. Dan Leonard
Code 8105
Naval Ocean Systems Center
San Diego, CA 92152 1

Dr. Alan F. Petty
Code 7930
Naval Research Laboratory
Washington, DC 20375 1

Dr. M. J. Fischer
Defense Communications Agency
Defense Communications Engineering Center
1860 Wiehle Avenue
Reston, VA 22090 1

Mr. Jim Gates
Code 9211
Fleet Material Support Office
U. S. Navy Supply Center
Mechanicsburg, PA 17055 1

Mr. Ted Tupper
Code M-311C
Military Sealift Command
Department of the Navy
Washington, DC 20390 1

	Copies	Copies
Mr. F. R. Del Priori Code 224 Operational Test and Evaluation Force (OPTEVFOR) Norfolk, VA 23511	1	

ROBUST DETECTION OF FADING NARROW - BAND SIGNALS

IN NON - GAUSSIAN NOISE

ERRATA

(1) Page 4, the end of the third paragraph should end with: VHSIC technology [16].

(2) Page 19, in the table heading, a diagonal separator is missing. The first entry should read: $\boxed{\epsilon \diagdown c}$.

(3) Page 25, second line before Eq.(3.10), (C.7) should be changed to (C.11).

(4) Page 32, Eq.(4.2) should be corrected to

$$\hat{A}_x(l)=Arg\{\sum_{i=1}^{n} s_i\, l(x_i - s_i \hat{A}_x)=0\} \quad , \quad \hat{A}_y(l)=Arg\{\sum_{i=1}^{n} s_i\, l(y_i - s_i \hat{A}_y)=0\} \qquad (4.2)$$

(5) Page 35, in Eq.(4.11), the definition of C should read: $C \triangleq \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} {s_i}^2$

(6) Page 44, line 9, *qalmost" should read: "almost".

(7) Page 61, line 5 from the end, should finish with: technology [16].